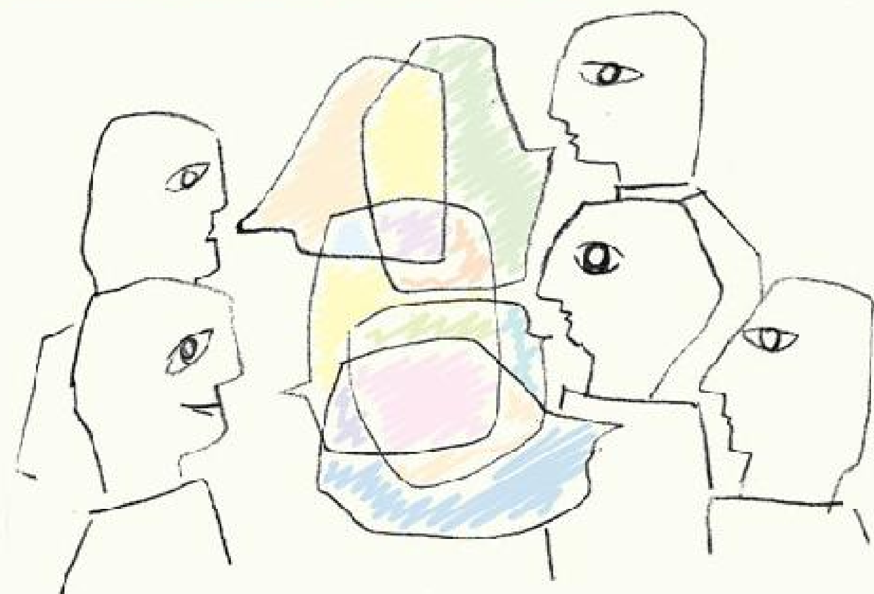
改变心理学的40项研究

7th Edition
第7版

【美】罗杰·R·霍克（Roger R. Hock）著　白学军 等 译

Forty Studies that Changed Psychology

Explorations into the History of Psychological Research



心理学史上影响无数人的、最重要的40项研究

20年来畅销不衰的心理学入门经典图书全新修订和升级

亚马逊心理学类畅销书Top100，亚马逊心理健康类畅销书No.1

中国人民大学出版社

VISIT...

LANZAROTE Caliente.com

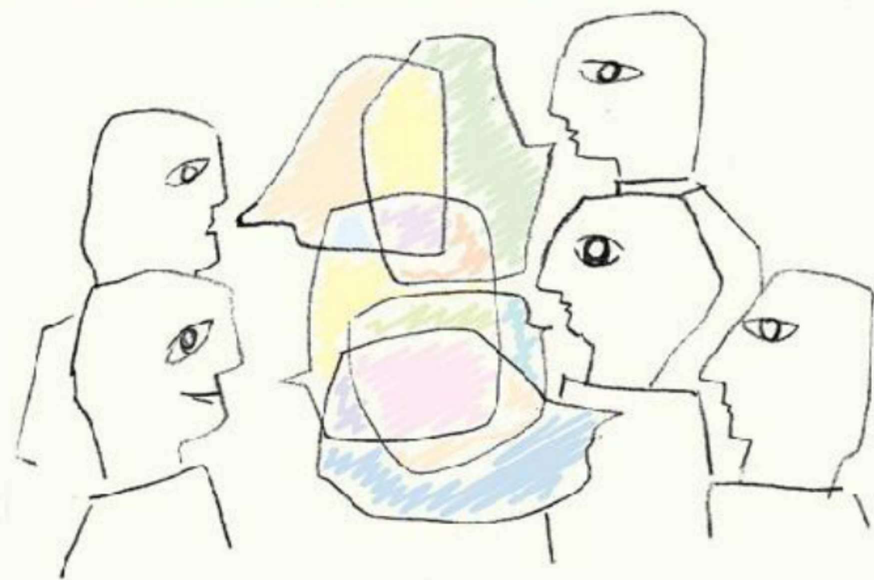
改变心理学的40项研究

7th Edition
第7版

【美】罗杰·R·霍克（Roger R. Hock）著　白学军 等 译

Forty Studies that Changed Psychology

Explorations into the History of Psychological Research



心理学史上影响无数人的、最重要的40项研究
20年来畅销不衰的心理学入门经典图书全新修订和升级
亚马逊心理学类畅销书Top100、亚马逊心理健康类畅销书No. 1

中国人民大学出版社

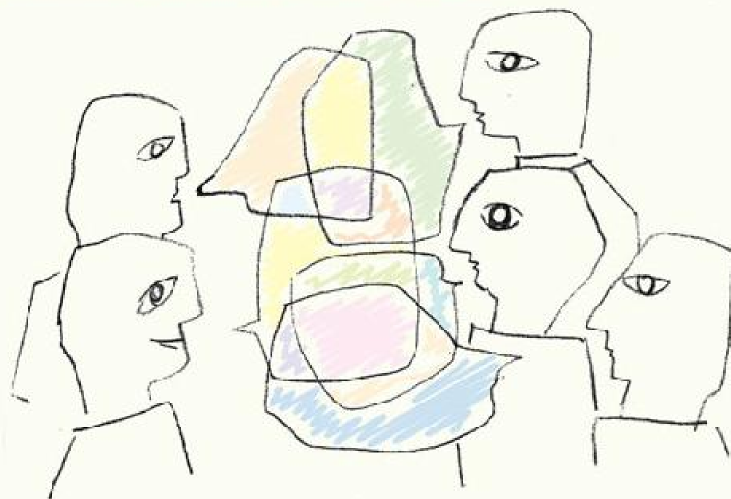
改变心理学的40项研究

7th Edition
第7版

【美】罗杰·R·霍克（Roger R. Hock）著　白学军 等 译

Forty Studies that Changed Psychology

Explorations into the History of Psychological Research



心理学史上影响无数人的、**最重要的40项研究**

20年来畅销不衰的心理学入门经典图书**全新修订和升级**

亚马逊心理学类畅销书Top100、亚马逊心理健康类畅销书No.1

中国人民大学出版社

版权信息

书名：改变心理学的40项研究

作者：（美）罗杰·R·霍克

出版社：中国人民大学出版社

出版日期：2015-08-01

ISBN：978-7-300-21668-3

价格：65.00元

目录
CONTENTS

译者序
前言
第一章 生物学与人类行为
研究1 一个脑还是两个脑
研究2 丰富的经历是否等于更大的大脑
研究3 人的本性是“天生的”吗
研究4 小心视崖
第二章 知觉与意识
研究5 注视偏好
研究6 睡眠，毫无疑问就会做梦
研究7 类别天成
研究8 行动，如同被催眠了一样
第三章 学习和条件反射
研究9 不只限于分泌唾液的狗
研究10 情绪化的小艾尔伯特
研究11 敲敲木头
研究12 观察到攻击行为……做出攻击行为
第四章 智力、认知和记忆
研究13 期望导致结果
研究14 你在哪方面更聪明
研究15 心中的地图
研究16 感谢记忆
第五章 人的发展
研究17 爱的发现
研究18 眼不见，不一定心不烦
研究19 你的品德如何
研究20 让你愉快的控制力
第六章 情绪和动机
研究21 性动机
研究22 我能读懂你的面部表情
研究23 观察你的情绪了吗
研究24 认知失调
第七章 人格
研究25 你能主宰自己的命运吗
研究26 男性化、女性化……还是双性化
研究27 和心脏赛跑
研究28 个人与集体

第八章 心理病理学
研究29 究竟谁是疯子
研究30 你将再次获得防御
研究31 习得性抑郁
研究32 拥挤导致行为失常
第九章 心理治疗
研究33 为自己挑选心理治疗师
研究34 缓解你的恐惧心理
研究35 投射出真正的你
研究36 编个故事吧
第十章 社会心理学
研究37 斯坦福监狱实验
研究38 从众的力量
研究39 你会伸出援手吗
研究40 无条件服从

译者序

白学军

教育部人文社会科学重点研究基地

天津师范大学心理与行为研究院

正如刻在德尔斐的阿波罗神庙的名言“认知你自己”所说，为了深化认识人类自身的心理问题，心理学家不断地进行探索，以期揭示人类心理的奥秘。

1879年，冯特创立心理学实验室，标志着科学心理学的诞生。从那时起，无数的心理学家从不同角度对人类心理问题进行了科学、系统的研究，其中的一些研究随着时间的推移变成了经典。经典的研究如同陈年老酒，越久越香！阅读经典总能让人受到启迪和教诲，也会给人以力量，或产生激励效果！

陶行知先生提出了好书的三个标准：（1）首先，要看这本书有没有引导人付诸行动的力量；（2）其次，要看这本书有没有发人深省的力量；（3）最后，要看这本书有没有追求新意和产生新价值的力量。本书中的每一项研究都符合以上三个标准。

2014年春夏之交，中国人民大学出版社的编辑联系我，告诉我他们获得了该书的最新翻译版权，希望我能够承担此项任务。介于此书是许多心理学爱好者、研究者、工作者非常喜欢的读物之一，因此，我欣然接受了这项光荣而又重要的工作，带领自己的学生们共同翻译此书。经过几个月的努力，现在终于可以交稿了。同时，看到自己的学生们也一天天地成长起来，我感到无比欣喜：“得天下英才而教育之，一乐也！”

子在川上曰：“逝者如斯夫。”从2004年我带领学生们第一次翻译此书，至今已超过十个年头了。我自己也从一名青年学者变成中年学者。一年又一年，时间怎么过得如此快呢？我常常情不自禁地感叹“时间都去哪儿了”。然而，许多美好的场景总是不断地激励鼓舞着我。记得许多次，心理学界的同仁给自己的学生们介绍“这位就是翻译《改变心理学的40项研究》的人”时，学生们都表现出激动、兴奋的神情，甚至有些学生还将自己新买的书拿来请我签名。另外还有些读者来信告诉我说“读了这本书后，我就喜爱上了心理学”；“读了这本书后，我才发现心理学是如此好玩”；“读了这本后，我才知道心理学如此有意思”；“心理学是科

学！读了这本书后，我也明白了如何从事心理学研究”，等等。

在全书的结构上，这一版没有变化，仍然由十个心理学专题构成，具体为：生理学与人类行为、知觉与意识、学习和条件反射、智力、认知和记忆、人的发展、情绪和动机、人格、心理病理学、心理治疗和社会心理学。每个专题包括4项研究。这十个专题涵盖了当今心理学的主流内容，受到的关注程度也非常高。

每一项研究包括题目、作者、时间、论文名称（或著作名称）、发表刊物（或出版社）、卷（期）和页码、问题提出的背景、理论假设、研究方法、结果、讨论、批评、最近应用和参考文献等内容，即一项完整研究的全部内容。通过阅读每一项研究，读者能够全方位地习得心理学研究的相关知识。

与本书的第五版相比，本版中更换了两项研究：第一项研究是20世纪60年代罗伯特·范茨用天才的方法，让人们知道刚出生的婴儿如何“认识”世界（从1961年起）的革命性发现；第二项研究是菲利浦·津巴尔多在20世纪70年代早期完成的“斯坦福囚犯研究”，主要探讨了在某种情境下权威和控制力对人类行为的影响。

除了上述变动外，作者对其他研究也作了部分修改，并根据最新的研究成果对其进行了丰富。

那么，我们如何才能在阅读本书后有所收获呢？

第一，读书贵有疑。明代学者陈献章说：“前辈谓学者有疑，小疑则小进，大疑则大进。疑者，觉悟之机也。一番觉悟，一番长进。”建议你读完每一项研究后，想一想，还有哪些问题值得进一步研究？还有哪些因素需要进一步控制？还可以将此研究拓展到哪些领域？

第二，反复阅读。陈寿的《三国志·魏志·董遇传》中有一句名言：“读书百遍而义自见。”对于经典研究，建议大家多读几遍，反复揣摩心理学大师们的智慧与创造。不仅要学习他们是如何发现问题和提出问题的，更要学习他们是如何巧妙地解决问题的，从而提高自己的“问题意识”和“解决问题的能力”。

第三，批判性阅读。在《孟子·尽心下》中有一句话：“尽信书，则不如无书。”首先，读者应仔细阅读每一项研究，尤其是重点阅读“批评”这部分内容，看一看前人是如何独立提出自己对经典研究的质疑的。在此基础上，尝试提出自己的想法。多次练习，就能提高自己独立思考的能力和创新能力。

参加本书翻译的人员：白学军（前言、第1章）、王永胜（第2章）、张慢慢（第3章、第4章）、李士一（第5章、第7章）、谭珂（第6章）、李馨（第8章）、刘丽（第9章、第10章）。翻译完初稿后，大家又进行了互校，然后由谭珂和李士一两位博士研究生对照原文仔细核

对了译稿。最后，我对全书进行了统校。

在本书前几版的翻译过程中，得到了著名心理学家、教育家杨治良教授和郭秀艳教授的鼎力相助，在此向他们表示衷心的感谢！

著名心理学家、教育家林崇德教授，把我领入心理学的殿堂。他用自己的智慧和勤奋取得了无数令人艳羡的成就，成为全国人民教师的楷模。能够成为他的学生，是我一生的骄傲和荣幸。林老师一直关心此书的翻译工作，我在此向他致以最诚挚的谢意！

著名心理学家、教育家沈德立先生生前对本书的翻译倾注了大量的心血。虽然他离我们而去了，但他留下的“爱国、尊师、勤奋、认真”的思想，激励着我们不断地前进！谨以此书献给沈先生，以表达弟子对他无尽的怀念！

父亲生前一直教导我：做好人、做好事。我一直坚守着父亲的忠告。在日常生活中，我经常是忙于工作而很少顾及家庭。一切家务全部由妻子承担，且得到岳父、岳母和母亲的大力支持与帮助，在此对他们表示感谢！

在翻译和校对书稿的过程中，虽然我们非常用心和认真，力争使翻译内容准确无误，但我深知自己学识修养还需要提高，因此，译文中如有不妥之处，敬请读者批评指正！

前言

欢迎大家阅读《改变心理学的40项研究》一书的第七版。20多年来，该书已成为世界各地许多大学和高中的主要参考书，且已被翻译为六种语言。第七版中的大部分内容在第一版时就已包含其中了，这表明那些标志性研究仍然对当今心理学的思想与研究有着深远的影响。这些研究和各版中增改的研究使我们得以一瞥科学心理学的诞生及发展，并使我们了解到人类在认识自身复杂天性方面所作出的努力。

在心理学的各分支领域中，对人类行为的众多研究都有着举足轻重、意义深远的影响。这些研究结果改变了我们对人类行为的认识，并为日后不计其数的实验设计和研究程序打下了坚实的基础。虽然有些早期研究取得的结果后来饱受争议和质疑，但它们在特定历史条件下的影响和作用却丝毫未被削弱。它们依旧为新的著述所引用，并成为学术探讨的话题，它们仍是教科书中的基本内容，并在心理学家的心目中占有独特的地位。

本书的构思源于我在心理学方面三十多年的教学经验。心理学教科书多以一些重要研究为基本内容，这些研究在相对短暂的心理学史上，对心理科学的形成有着举足轻重的作用。然而，教科书通常简练概括这些研究的过程和结果，却很少对原创性的重要研究给予应有的关注，从而大大削弱了科学探索的生命力和兴奋感。有时，对这些研究方法和成果的介绍方式甚至会影响读者正确理解研究过程及真正意义。这绝不是为了批评教科书的作者们，因为教科书的篇幅有限，他们不得不对研究的内容和细节进行艰难的取舍。然而，出现这种情况是令人遗憾的，因为现代心理学的基础是科学研究。在一个多世纪里，有许多构思独特、设计精巧的研究，不断地拓展和完善着我们对人类行为的认识与理解，使之达到今日如此深入的程度。

本书试图填补心理学教科书与心理学研究之间的鸿沟，它也是对心理学史上重大事件的一次回顾。我希望通过自己的介绍能使这40项研究得到真实再现，以让读者亲身感受它们。本书谨献给所有希望深入了解心理学真正起源的人。

研究的选择

本书所涵盖的研究是根据心理学教科书、学术期刊和许多心理学分支学科的权威专家的建议精心挑选出来的。虽然对应选多少研究事先并无计划，但是从历史的角度和本书的篇幅两方面来考虑，选择40项研究似乎恰到好处。所挑选的这些研究，在心理学史上也许是最著名、最重要或是最有影响的。我用“也许”这个词是因为本书的许多读者可能会对其中的一些选择持不同的意见。毫无疑问，要列出一张让所有人都满意的40项研究报告的目录是不可能的。然而，本书所包含的这些研究一直被频繁地引用，在它们发表以后，引发了激烈的争论，引出了很多的相关研究，开创了心理学探索的新领域，极大地改变了我们对人类行为的认识。这些研究是按其所属的心理学分支组织起来的，主要包括：生物学与人类行为；知觉和意识；学习和条件反射；智力、认知和记忆；人类的发展；情绪和动机；人格；心理病理学；心理治疗学和社会心理学。

本书的体例

这本书不可能将原始研究报告全部包括在内。为了便于读者更好地理解每一项研究，本书采用了统一的格式对这些原创研究进行了分析和概括，每项研究都包含以下内容：

1.一篇确切的、方便查找的关于该原始研究出处的参考文献；

2.该研究背景的简要介绍及开展此项研究的原因；

3.该研究检验的理论命题或假设；

4.对实验设计和方法的详细介绍，包括被试的选择和组织、仪器和材料的使用、进行研究的实际步骤；

5.用清晰的、易理解的、非技术性的、非统计术语的通俗语言总结研究结果；

6.根据研究者在原文中对结果的讨论来解释该项研究发现的意义；

7.该研究在心理学领域内的重要性；

8.简要讨论后续研究的支持性或反对性结果，以及本领域其他研究对该研究提出的批评与质疑；

9.介绍一个近期在其他文章中应用和引用此研究的例子，以说明该研究的持续影响力；

10.与本研究有关的进一步的和最新的参考资料。

通常，要想理解科学家所用的语言并不容易（甚至对其他科学家而言也是如此）。本书的主要目标是使这些研究变得更有意义，从而更容易被读者所接受，并使读者能够体验到这些引人注目的重要发现中令人激动的美妙之处。我尽可能对这些研究进行适当的简化论述，使之易读易懂，同时又注意保持原有研究的精华部分，以便读者明了其重要影响。

第七版中的新内容

《改变心理学的40项研究》第七版中包含了许多重要的实质性变化，新增了两项关于人类如何知觉世界的研究报告，它们在心理学史上影响非凡：第一项是1973年埃莉诺·罗施（Eleanor Rosch）的革命性工作，她通过巧妙的方法来研究人类大脑对颜色的分类；第二项是近年来的研究，它使用高技术设备（主要是MRI）直观地揭示出人脑活动。

此外，本书还更新了多篇研究报告中的“近期应用”部分，反映了进入21世纪以来，这40项研究被大量引用的情况。每年，本书所提及的40项研究被引用次数超过1000次！书中选取了其中一些研究作简要介绍，以便读者能够去亲身体验这40项研究对心理学的持续影响。在每一章结束时，新的研究将连同其他相关的资料一起列出。读完这些后，读者将体会到这些研究对心理学发展所作出的巨大贡献。

从第六版正式出版后的几年以来，我继续与许多心理学分支学科的专家进行交谈，共同探讨了新一版中可能出现的变化，从中获得了许多有益的意见和建议。这些研究同行时常会提到两项颇具影响力的研究，我也曾一度考虑过它们，所以在这一版中我把它们加了进来。两项新增加的研究以其独特的视角，拓宽了我们对人类本性中非常基础的两个方面的认识，增加了我们对人类体验复杂性和多样性的理解。

在第七版中新增加的第一项研究，实际上在20多年前第一版中已经包括，这次是重新选入。在过去的岁月里，许多学者告诉我，这项研究极为重要，不应该从本书中拿掉。这项研究就是由埃莉诺·罗施所完成的，她提供的研究证据表明，我们天生就能对颜色进行分类，许多颜色是人们并未学习过的，而是通过预存的颜色原型对它们进行分类，这种预存的颜色原型程序化地存在于大脑之中，并且通过进化被一代又一代地传递着。当我们认识到颜色的再认能力对早期人类的生存和物种存活具有非常重要的作用时，我们就能理解该研究的意义所在。虽然早期人类还不能命名颜色，但是颜色对于生存的价值已经发挥作用了，如“上次我吃了黄色的叶子，使我大病一场，所以我必须远离它。如果我再吃它，我可能会中毒而死”。

在第七版中新增加的第二个研究与其他研究稍有不同，与其他研究相比，该项研究比较新（2003年）。然而，之所以将此项研究收录，是

因为科学研究的主流已转到对大脑的认识，而且科学界已能揭示大脑是如何工作的。一种新的范式已经允许研究者"观察"人们在进行心理活动时的大脑功能。它还有可能使我们完全能"看到"人们正想什么！允许我们进行上述活动的工具就是核磁共振（MRI）。虽然MRI已经不是一项十分新颖的技术了，但是研究者现在才用它来观察人们思考时的大脑功能。这被称为功能性核磁共振或fMRI。MRI的这一科研用途仍然处于发展阶段，且人们对其最大潜能仍存有争议。但是当你阅读该项研究时，你就会见证fMRI那令人惊叹的用途。

所有40项研究，无论其发表的早晚，我们随后都对一个共性问题进行了讨论，这个共性问题就是研究的伦理性。作为科学的心理学，一个最重要的基石是心理学家在以人或动物为被试开展研究时，必须遵守严格的职业道德规范。下面让我们简要总结一下社会科学家在开展研究工作时所遵循的伦理原则。

以人或动物为研究对象的道德问题

没有研究对象，就没有科学研究。物理学的研究对象是物质和能量；植物学的研究对象是植物；化学的研究对象是分子、原子和亚原子粒子；而心理学的研究对象则是人。有时，一些研究中所用的特定程序或行为不能用于人，于是就用动物来替代。然而，以动物为对象的研究目的是为了更好地了解人类，而不是动物本身。在接下来的文章中，读者将读到一些以人和动物为对象的研究。有些可能会引发人们对研究中的道德问题的思索。

如果所介绍的研究中存在有让人痛苦或者紧张的研究程序时，我们会提醒大家注意研究中的道德问题。但由于这也是一个敏感的话题，我在这里就简要谈一下当代心理学家所遵循的伦理原则，就当为本书后面要介绍的研究内容作铺垫。

以人为被试的研究

美国心理学会（The American Psychological Association，APA）制定了严格、明确的准则，用来指导以人作为被试的心理学实验研究。下面是这些准则的部分内容：

> "心理学家必须为被试的利益着想，谨防对他们造成伤害。在职业活动中，心理学家要尽力保护研究对象的利益和权力……当职责或关注的问题之间发生冲突时，心理学家力争以高度负责的方式解决冲突，努力避免伤害或者使伤害降低到最小……心理学家必须坚持行为的职业准则，明确他们的专业角色和职责，为他们的行为承担相应的责任，并致力于妥善处置可能导致侵犯或伤害的利益冲突……心理学家必须尊重全人类的尊严和利益，尊重个体的隐私权和自主权。"[摘自《心理学家的伦理原则和行为法规》（Ethical Principles of Psychologist and Code of Conduct, 2003）；见www.apa.org/ethics/]。

今天，根据这些准则，在以人为被试的研究中，必须遵守以下几条原则。

1.免于伤害

这对大家来说是不言自明的。研究者有责任保护被试免于伤害，难道还有其他情况吗？答案是确实有！但这并不是一个一成不变的规则。本书中的一些研究同时讨论了实验是否侵犯到被试的权利，以及研究者是否遵循了相应的伦理原则。需要指出的是，保护被试免于伤害必须扩展到实验之外，如果被试在实验后感到不安，他们可以与研究者联系，一起商讨问题所在。

2.知情同意

研究者必须事先告知被试实验的目的和程序，以便让被试能够做出是否参加实验的决定。如果某人同意参加实验，就称为知情同意（informed consent）。在本书中，读者可能会看到，有时实验的目的不能事先告诉被试，否则将会使被试的行为发生变化并污染实验的结果。在这种情况下，研究者就要隐瞒实验目的。即使如此，研究者仍应尽量给被试足够的信息，做到"知情同意"。并且，实验所隐瞒的部分必须合理地建立在潜在的研究结果重要性之上，在实验结束后，应将相关内容透露给被试知晓，做到"知情同意"。涉及儿童和青少年的研究，以上伦理原则同样适用，且其家长或监护人需签署知情同意书。

3.随时退出实验的自由

在以人为被试的研究中，被试应该知道在整个实验过程中，他们随时有退出实验的自由。这似乎是一条不必要的规则，因为很显然，当被试感到实验程序让他们十分不适时，就能简单地选择退出。然而，事情并不总是如此简单。例如，在大学里，许多心理学实验的被试是那些为得到学分而参与实验的学生。他们会认为，中途退出实验会影响他们得到学分，因此他们并不自由。付费被试在参加心理学实验时，如果研究者让被试觉得只有做完实验才能获得报酬，那么这种不道德的诱导会使被试失去退出实验的自由。为了避免上述问题的出现，被试应在实验开始前，仅凭他们的参与便得到相应的学分或报酬。

4.保密

除非得到被试同意，否则所有的实验数据都应保密。尽管这并不意味着结果不能公布或发表，但是其公开的前提条件是必须隐匿任何与个人身份有关的数据信息。通常，研究者甚至根本不采集有关被试身份的信息，而是将所有被试的数据合并，计算其群体的平均差异。

5.执行任务

绝大多数心理学研究（包括所用方法）在进行时或完成后对被试没有任何伤害。然而，即使看起来无害的实验程序有时也可能使被试产生沮丧、尴尬、担心等负面影响。防止这种情况的常用措施是询问被试执行任务的情况。在被试完成一个实验后，特别是那些含有任何形式的欺骗的实验，被试应被告知真相。在此过程中，实验人员应该给他们说明实验的真实目的和目标，而且被试有权对有关问题进行询问。如果实验有可能对被试造成一定的影响，那么研究者应该将自己的联系方式留给被试，必要时作进一步讨论。

读者在阅读本书时可能会发现有些研究似乎违反了伦理准则。需要指出的是，这些研究都是在制定正式的心理学研究伦理原则之前完成的，现在已经无法重现这样的研究了。但是伦理原则的缺乏并不能成为早期研究者违背道德的理由，对他们的判断应该由我们每个人自己来完成。同时，我们还应像众多心理学家所做的那样，从过去的失误中吸取教训。

以动物为被试的研究

在科学界内外讨论得最热烈的话题之一就是动物研究中的道德问题。随着动物保护组织数量的不断增加，他们的声势也不断壮大。今天，之所以出现有关动物被试的争论比有关人类被试的多，主要原因可能是动物不像人类那样能够获得知情同意和自由退出实验等准则的保护。另外，一些激进的动物保护主义者认为，所有的生命都是有价值的，它们能感知疼痛。根据这一观点，动物和人类具有同样的价值。任何一种对动物的利用形式都是不道德的。这种利用形式包括食用鸡肉、穿着皮革和饲养宠物等（按照某些动物保护主义的观点，宠物是奴隶的一种形式）。

争论双方中的一方认为，用动物做研究是不人道、不道德的，是应被禁止的。然而，几乎所有的科学家和大多数美国人认为，在科学研究中有限制地、人道地使用动物是必须而有益的。许多挽救生命的药物和医疗技术都是以动物为被试发展起来的。在心理学研究中，常用动物来研究抑郁、大脑的发育、过度拥挤和学习过程等问题。其主要原因在于，在人身上从事类似的研究明显是不道德的。例如，你想知道玩具与活动的丰富与否对婴儿大脑发育和智力的影响，便把婴儿安排在这两种条件下进行实验，这种做法显然是不可行的。然而，大多数人则会同意用老鼠来进行研究，这既能获得对人类而言非常重要的研究发现，同时又避免了道德问题（见本书研究2中Rosenzweig & Bennett的研究）。

美国心理学会除制定了以人为被试的研究所应遵循的原则外，还制定了严格的规则来规范以动物为被试的研究。这些规则包括让研究者为动物提供适当的居住空间、合理的喂养、清洁的环境和健康保护，禁止

对动物施加所有不必要的伤害。美国心理学会的《关心和使用动物的道德行为准则》（2004）的部分内容如下：

> 动物在实验期间应给予人道主义关怀和健康的生存条件……鼓励心理学家为实验动物创造丰富的环境，心理学家应该掌握如何保证实验动物健康以及营造丰富生活环境的最新文献资料……当行为程序可以选择时，应该使用对动物造成最小伤害的方式。当使用厌恶性条件时，心理学家应该调整刺激的参数，使其既符合研究的目的，又可以把对动物的伤害降低到最小。只要合乎情理的，应鼓励心理学家亲自尝试引起痛苦的实验刺激（见//www.apa.org/science/anguide.html）。

本书中有几项以动物为被试的研究。除必须考虑这些研究的道德问题外，动物研究的结果也很难直接推论到人类。我将在涉及以动物为被试的研究的各章节中讨论这些问题。无论是心理学研究者，还是学习心理学的学生，每个人对动物研究都应有自己的观点，并在具体研究中决定使用动物作被试是否适合。如果你认为在某些情况下用动物做实验是可接受的话，那么对于本书中以动物为被试的研究来说，你有必要对这些研究结果作出评判，看它的价值是否足以支持它所用的方法。

最后需要关注与动物研究有关的一些工作进展，即公众关心的虐待动物问题已经引起有关部门的关注。马萨诸塞州的剑桥城拥有哈佛大学和麻省理工学院等机构，是重要的国际研究中心之一，那里的健康与医疗系（见http://www.cambridgepublichealth.org/services/regulatory-activities/lab-animals/lab-animals-overview.php）设立了一个叫作动物实验委员会的专员职位。这是第一个该类型的管理职位，由朱莉·梅德利（Julie Medley）博士担任。目前剑桥城内有44家研究机构，饲养着约20万只动物。该专员负责监管研究过程中的每一环节对动物是否人道，包括从动物生活方面到研究计划中使用动物的方法。如果在某实验室中发现有违反剑桥城法律的行为（即对待动物不人道），该专员有权对该实验室处以每天300美元的罚款。在写本书的过程中，仅有一家因违规而被罚款，该机构因故意无视动物处理法而被罚了总数40000美元（共计有133天违法）。在所有其他情况下，一旦发现违反了规定的设施，实验室会自愿快速地作出修改调整。

在本书中，读者将要看到的这些研究，已经以不同方式在不同程度上为全人类作出了贡献。心理学研究的历史虽然相对较为短暂，但是它却充满了丰富多彩、激动人心的对人类本性的发现。

感谢

谨向出版人Charlyce Jones Owen女士致以诚挚的谢意，本书从第一版伊始就得到了她的大力支持。我非常感谢培生公司心理学部总编Jessica Mosher女士对此项目的大力支持和源源不断的、富有才华的帮助。

对于此版书，我要衷心感谢培生公司的执行心理学总编Stephen Frail、项目发展总编Maddy Schricker和第七版GEX出版服务项目经理Michelle Durgerian。我以个人名义对才华横溢的Bruce Kenselaar和Suzanne Behnke为本书过去几版和这一版的封面设计表示感谢。

我真诚地感谢心理学同行为本书过去几版花费了他们宝贵的时间、精力，并与我交换了他们的意见、建议。我试图利用每一次机会将他们有价值的见解包含在本书中。

最后，我要向过去20多年里，直接或间接给予过本书帮助的所有家人、朋友和我的学生们，表达我深深的祝福和最衷心的感谢。

罗杰·R·霍克

第一章 生物学与人类行为

几乎所有基础心理学教科书的开篇都与研究人类行为的生物学有关。这并不能简单地归因为惯例，而是因为生理过程构成了所有行为的基础。心理学的任何一个分支都是建立在生物学基础上的。心理学中研究这一过程的分支称作“生理心理学”（physiological psychology）或“生物心理学”（biological psychology），重点研究大脑和神经系统的相互作用，研究人类通过感官从外界环境获取刺激和信息的过程，研究大脑通过组织这些信息形成人类对客观世界的知觉方式。

作为心理学的这一基本组成部分的代表，本章选录的研究涵盖范围很大，也是影响力最大、引证次数最多的研究：第一项研究报告探讨了有关大脑左右两个半球功能特异性的著名研究，它使我们现有的有关大脑功能的知识得以形成；第二项研究报告证明了幼年时期的刺激将有可能促进大脑的高度发展，这一发现让科学界感到震惊；第三项研究报告代表了很多心理学家在思考对人类行为、人格、社会交往产生的基本原因时所发生的重要变化，即关于基因重要性的一种新观点；第四项研究报告用于研究婴儿深度知觉的著名方法“视崖”装置的发明。所有这些研究报告，尤其是后两项，提出了一个贯穿心理学各领域的基本议题，并引发了连续不断、引人入胜的讨论：先天—后天之争。

研究1 一个脑还是两个脑

Gazzaniga, M.S.(1967).The split brain in man.Scientific American, 217(2), 24-29.

也许，你已经意识到自己大脑的两半球并不相同，且它们的功能也不同。例如，大脑的左半球负责身体右侧的行动，反之亦然。不仅如此，大脑的两半球在其他许多方面表现出更大程度的功能定位。

大脑左半球控制语言能力，而大脑右半球则更多地涉及空间关系，如艺术活动所需要的相关能力。对大多数人来说，这些知识已经相当普及了。众所周知，因中风或遭遇意外事故的受害者，如果其大脑的左半球受到损伤，通常会使他们丧失语言能力（通过训练，这种能力一般可以恢复）。许多人相信，大脑的每侧或每个“半球”确实是可以完全分离的智力系统，它们各自具有独立的学习、记忆、认知世界甚至感受情绪的能力。这一共识的建立是多年来认真研究割裂脑的结果。

这方面研究的先驱是罗杰·W·斯佩里（Roger W.Sperry，1913年~1994年），他的研究在本章涉及的论文发表前15年就开始了。在以动物为被试的早期研究中，斯佩里获得了许多重大发现。例如，通过外科手术切断一只猫大脑两半球之间的联系，并改变它的视觉神经，使猫的左眼只传输信息至左半球，右眼只传输信息至右半球。手术后，这只猫表现得很正常，并且几乎没有显示出不良后果。然后蒙住它的右眼，让它学习一种新的行为，比如穿过一个小迷宫找到食物。在这只猫能熟练地通过迷宫后，放开右眼，再蒙住它的左眼。然后把它再放入迷宫，它的左脑对在哪儿转弯根本没有印象，不得不重新开始熟悉整个迷宫。

斯佩里在此后的30多年里进行了许多相关研究，并且凭借在大脑两半球功能定位方面的成就而在1981年获得了诺贝尔奖。20世纪60年代初，他的研究方向转向以人为被试，在加州理工学院，迈克尔·加扎尼加（Michael Gazzaniga）加入了他的研究团队。虽然斯佩里被认为是割裂脑研究的奠基人，但本篇选择了加扎尼加的文章，这是因为这篇文章清晰、简洁地概括了他们早期以人为被试的合作研究，并且一直被许多普通心理学教科书引用。这一选择绝不是忽视或掩盖斯佩里在此领域的领导地位或伟大贡献。在很大程度上，加扎尼加把他在大脑半球功能定位方面的早期研究及发现归功于斯佩里（见Sperry，1968；Puente，1995）。

为了理解割裂脑研究，首先需要介绍一些人类生理学方面的知识。大脑的两半球通过胼胝体（corpus callosum）相互联系，胼胝体是由约2亿根神经纤维构成（见图1—1所示）。如果胼胝体被切开，大脑两半球连接的主要途径就被切断，那么两半球的功能就彼此独立了。所以，

如果我们想要单独研究大脑的每一个半球，只需要通过外科手术切开胼胝体即可。



图1—1 人类大脑的胼胝体

资料来源：引自3D4医学/图像研究者公司。

但是，科学家能把人的大脑分开吗？这听起来有点像是科学怪人弗兰肯斯坦博士的故事[1]。显然，伦理道德决不允许单单为了研究大脑两半球的特定功能就对人采取这样残忍的方式。然而，20世纪50年代后期，医学界给了心理学家一个黄金般的机遇。在一些非常特殊的极端案例中，对无法控制的癫痫患者来说，切断胼胝体可以极大地减少、甚至消除癫痫发作。对于那些没有其他方法可以救治的病人，这种手术作为最后的治疗方法，在那时是（并且现在也是）十分成功的。当这篇文章在1966年完成时，已经进行了10例这样的手术，其中4名患者同意参加斯佩里和加扎尼加的测试，以确定这种外科手术对他们的知觉能力和智力技能产生了什么样的影响。

理论假设

研究者力图揭示人脑两半球能独立行使职能的范围以及它们是否有互相分离的、独一无二的能力。如果大脑两半球之间的信息传输被中断，你身体的右侧会不会突然变得无法与左侧协调？如果语言由大脑左半球控制，在这个手术后，你说话的能力和对词语的理解能力会受到怎样的影响？思维和推理过程是否都独立存在于两个半球中？如果大脑真的被分成两个脑，当这两个脑不再交流信息时，一个人还能正常地行动吗？既然感觉输入是同时作用于左右两侧的，那么视觉、听觉、触觉等会受到怎样的影响？在斯佩里和加扎尼加的割裂脑研究中，他们试图回答上述问题。

方法

研究者设计了三组不同类型的测试用来揭示患者的一系列心理（认知）能力。一种测试是检测视觉的。他们设计了一种方法，使一张印有一个物体、一个单词或者只言片语的图片只能投射到大脑的右半球或左半球的单侧视觉区域（称作“视野”），而不是同时投射到大脑的两个半球。应该注意的是，正常情况下，两只眼睛能分别向大脑的两半球同时传递信息。然而，当物体或单词出现在你面前的某一精确位置，你的视线固定在一个特定的点上时，你大脑的左视野或右视野就可能单独获得图像。

他们设计的另一种测试是有关触觉（接触）刺激的。患者只能去触摸，却看不到给出的物体、字母块或由字母块组成的单词。这种装置上有一个底部带有窟窿的屏幕，以便被试能摸到物体，却看不见所摸的物体。视觉和触觉装置可以同时使用。例如，当一张画有钢笔的图片被投射到大脑的一个半球时，将钢笔及一些其他物品放在屏幕后，以便被试触摸（参见图1—2）。

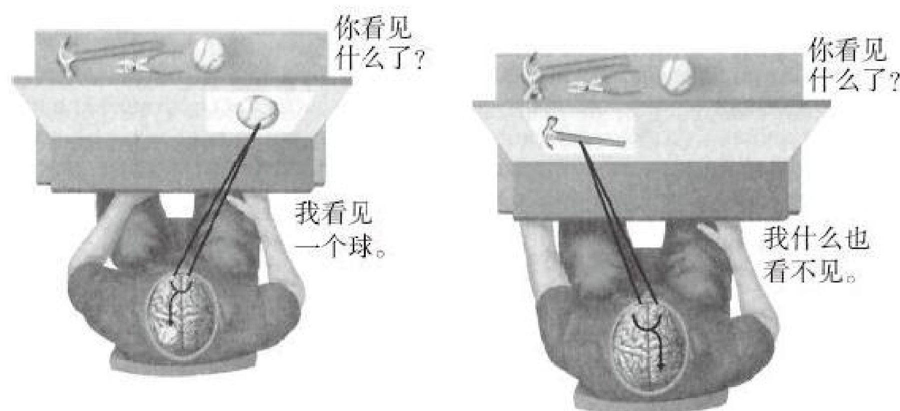


图1—2 对割裂脑患者进行视觉测验的典型装置

最后是测试听觉能力，这似乎需要更多的技巧。当声音传到你的一

只耳朵时，感觉信息同时输入你大脑的两半球。因此，想要限制听觉信息只输入大脑的一个半球，即使对割裂脑患者，也是不可能的。然而，限制大脑一个半球对这些输入信息作出反应则是可能的。做法如下：想象有几种普通的物品（一把汤勺、一支钢笔、一个玻璃球）放在布袋里，然后要求你摸出特定的物品。你也许毫不费力就能做到。如果你把左手放进袋子里，这只手是由大脑的右侧控制的，反之亦然。你认为大脑的一个半球可以单独完成这个任务吗？就像你在后面将会看到的，大脑的两个半球在执行这项听觉任务时并没有同等的能力。如果不是要求你找到特定的物品，而只是要求你将手伸入袋子，通过触觉来识别物品，那又会怎么样呢？这对于你并不困难，而对于割裂脑患者来说则相当困难。

加扎尼加综合上述这些测试方法，揭示出大脑功能之谜。

结果

首先需要指出的是，你应该知道在实施这种极端的脑部外科手术后，患者的智力水平、个性特点、特有的情绪反应等相对来说并没有改变。他们显得非常愉快，感到一种摆脱病魔的自由。加扎尼加报告说，一个手术后很虚弱的患者还开玩笑说，他感到“分裂的头痛”。然而，当测试开始时，这些被试却表现出许多不寻常的心理能力。

视觉能力

第一个测试中，用带有一排水平排列的灯泡的木板。患者坐在木板前，盯着中间的一个灯泡看时，这些灯泡将在其左、右视野依次闪烁。然而，当研究者要求病人报告他们看到了什么时，他们却说只看到板上右边的灯闪烁过。之后，研究者只让处于患者左视野上的灯泡闪烁，病人报告说他们什么也看不见。对于这一现象，唯一符合逻辑的结论是大脑的右半球是一个盲区。随后一件奇怪的事情发生了。灯泡再一次闪烁，这次要求患者指出灯泡闪烁的位置。尽管他们说只看到右边灯泡的闪烁，但他们能指出整个视野中的灯泡。使用了这种指认的方法后，我们发现大脑的两半球都能看见灯泡，并且在视知觉方面能力相当。在这里，重要的一点是病人不能说出他看到了全部灯泡，并不是因为他们不能看见，而是因为语言中心位于大脑的左半球。换言之，为了能说出你看见的东西，就不得不让你大脑的左半球看见这一物体。

触觉能力

你可以自己进行一下这个测试。把你的双手放到身后，然后让别人把熟悉的物品（一把勺子、一支钢笔、一本书、一只手表）放在你的右

手或左手里，看你是否能辨别它们。你可能觉得这样做不太难，是不是？这基本上就是斯佩里和加扎尼加对割裂脑患者所做的测试。用上述方法把某一种物品放在患者的右手中，不让患者看到或听到，物品的有关信息只传递到他们大脑的左半球，在这种条件下，患者能够叫出物品的名字，并能描述它，指出它的用途。然后，当同样的物品被放在患者的左手（与右半球相关联）时，患者不能说出物品的名字，也不能描述它。但患者知道物品是什么吗？研究者为了找到答案，要求被试在他面前的各种物品中找出与左手中的物品（记住，被试并没有看见它）相配的物品。他们就像你我一样，很容易地将它找出来。这再一次表明，口头表达能力位于大脑的左半球。记住，你能说出放在左手里的物品的名字是因为信息从你大脑的右半球经由胼胝体传到了左半球，你的语言中枢在那里会说“那是一把勺子”。

视觉与触觉联合测试

综上所述，这两类测试不仅支持了上述发现，还得出了额外的、有趣的结果。如果只向被试大脑的右半球呈现某一物体的图片，那他们既不能说出物品的名字，也不能描述它。事实上，患者也许根本没有言语的反应，或否认有任何东西出现。但是如果允许患者伸出左手（仍然只用右半球）到屏幕下触摸物品，他们总能找到与视觉呈现相同的物品。

测试发现，右半球也能很好地思考和分析物体。加扎尼加曾报告说，在展示某一物品的图片（比如香烟）给被试的右半球时，允许被试在屏幕后触摸10种物品，但里面却没有香烟，并让其选择最接近图片的一件物品——在本研究中是一个烟灰缸。加扎尼加继而解释道：

> “然而，太奇怪了！即使在他们正确反应之后，并在他们左手握着烟灰缸时，他们也不能说出或描述实物烟灰缸以及图片上的香烟。显然，在知觉和知识方面，左半球和右半球完全分开了。”（P.26）

另外一些测试结果很清楚地表明右半球具有语言加工能力。有这样一个著名的、极富创造力和启示性的视觉测试装置。当“HEART”这个词被投射到患者的视野中，其中右视野可以看到“HE”，左视野可以看到“ART”。现在，根据大脑两半球的功能，你认为患者将会口头报告看到了什么？如果你说是“ART”，那么你答对了。然而，这一实验的启示性在于，当被试面对印有单词“HE”和“ART”的两张卡片，并且要他们用左手指出所看见的单词时，他们全都指向“HE”！这证明大脑右半球能理解语言，虽然它使用了不同于大脑左半球的方式，即一种非言语的方式。

对患者进行的听力测试也得出了相似的结论。要求患者把左手伸入一个他们看不到的混杂着不同物品的袋子里，让他们拿出某一特定的物品（一块手表、一个玻璃球、一把梳子、一枚硬币），他们没有丝毫困难。这表明大脑右半球正在理解语言。甚至研究者只对目标物品的某一特征进行描述，要求患者去寻找时，他们仍能做到准确无误。加扎尼加

举了一个例子：当要求患者在一个装满不同塑料水果的袋子里去找出“猴子最喜欢的水果”时，他们拿出一个香蕉；当要求患者找出“新奇士（Sunkist）销售很多的一种水果”时，他们拿出一个橙子。然而，如果同样的水果放在患者的左手里而不让他看见时，他们就不能说出那是什么。换言之，当要求被试用言语作反应时，大脑右半球则无法执行。

大脑两半球具有惊人的不同之处，我们的最后一个例子采用了放在屏幕后桌子上的塑料字母块。主试要求患者用左手将这些字母块拼成各种单词时，他们很容易地做到了。即使把能拼成特定单词的三四个字母块放在屏幕后面，他们也能用左手依照单词正确地排列字母。然而，完成这一任务后，这些患者却不能说出刚刚拼写的单词。很明显，大脑的左半球在言语上优于右半球（这在一些左利手人中可能正好相反）。那么，大脑右半球有什么优势技能呢？斯佩里和加扎尼加在这一早期研究中发现，在涉及空间关系和形状的视觉任务中，左手的成绩更好（即使这些患者都是右利手）。就像在图1—3中看到的，复制三维图画（用铅笔在屏幕后复制图画），左手的成绩更好。

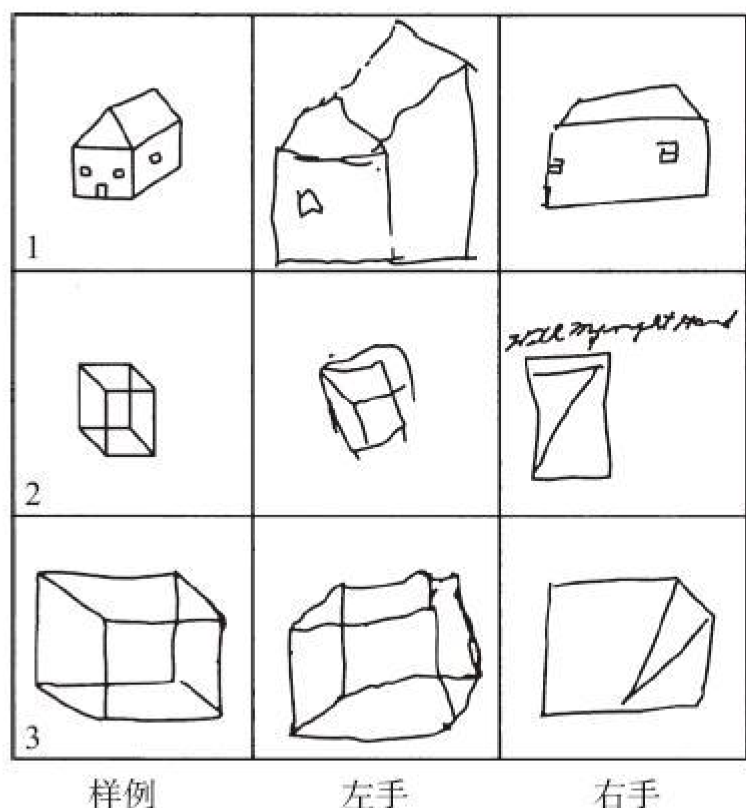


图1—3 割裂脑患者所画的图

资料来源：引自Michael S.Gazzaniga“割裂脑人”。

最后，研究者想要揭示割裂脑患者的情绪反应。进行视觉实验时，斯佩里和加扎尼加突然给被试的左半球或右半球呈现一张女人的裸体照片。例如，他们把这张照片投射到一名女性患者大脑的左半球。

“她大笑，并且能用语言说出这是张裸体照片。之后当这张照片呈现在她大脑的右半球时，她说……她什么也没看见，但她脸上立即闪过狡黠的微笑，并开始咯咯地笑起来。问她笑什么时，她说：‘我不知道……没什么……

哦——那个机器真滑稽。’虽然大脑右半球不能描绘所看到的东西，但视觉刺激仍能在大脑右半球引起情绪反应，就像在大脑左半球引起的一样。”（P.29）

讨论

从这篇研究报告得出的总结性的结论是，每个人的颅骨内有两个不同的脑，每个脑都具有复杂的能力。加扎尼加提出了一种可能性，即如果我们的大脑真的是两个脑，那么当大脑的两半被分开后，也许我们就有双倍处理信息的潜能。的确，有些研究证据表明，在正常人完成一项认知任务的时间内，割裂脑患者能完成两项任务。

结果的意义

斯佩里、加扎尼加的这些研究结果及后续研究在许多领域都是极其重要并影响深远的。现在我们已经知道，大脑的两半球有许多特有的技能和功能。大脑左半球更擅长于言语、写作、数字运算、阅读，等等，是主要的语言中枢。大脑右半球则在加工人像、解决空间关系问题、符号推理、艺术活动等方面更有优势。

随着对大脑两半球功能定位的认识的深入，我们可以有效地治疗中风或脑部损伤的患者。明确损伤的位置，就可以预测当病人康复后会存在哪些后遗症。通过这些知识，可以制订出恰当的重新学习、重新适应的计划，用来帮助病人尽可能快速、全面地恢复健康。

加扎尼加和斯佩里在这一领域进行了多年的研究以后，得出了“大脑的每个半球的确有各自的思想”的结论。在后来的研究中，他们对割裂脑患者测试了比上述问题更复杂的问题。其中一个是：“你会选择什么职业？”一名男性患者口头（大脑左半球）回答说，他想当个绘图员，而他的左手（大脑右半球）通过触摸大写字母拼出了“automobile race（赛车）”（Gazzaniga & LeDoux，1978）。事实上，加扎尼加向前推进了他的理论。目前他坚持认为，即使大脑完好无损的正常人，大脑两半球之间的联系也并不完全充分（Gazzaniga，1985）。例如，如果特定的信息，比如形成某种情绪的信息，没有以语言的方式存储起来，大脑左半球就不可能提取它。结果就可能是你感到伤心，但却说不出原因。由

于这是一种令人不适的认知困境，因此大脑左半球会试图用语言来解释这种悲伤（毕竟，语言是它的主要工作）。然而，由于大脑左半球并不具备所需的充足资料，所以它的解释可能完全是错误的！

批评

多年来，很少有人对斯佩里、加扎尼加及其他人在割裂脑研究中得出的结果产生质疑。有关这项研究的批评转而集中在“大脑左、右半球功能特异性”这一观点向大众文化及传媒传达的方式上。

现在，有一种说法被人们广泛接受，即有些人的大脑右半球或左半球更为发达，或是为了提高你的某项技能，你需要发展大脑的某半球。美国芝加哥大学的生物心理学家杰瑞·利维（Jerre Levy）一直在努力消除有关人具有两个功能独立的大脑的观念，并在这方面处于最前沿。她认为，正因为大脑两半球有不同的功能，所以它们的功能必须整合，而不是如大多数人所相信的那样，需要互相分离。通过整合，大脑能用超过并不同于任何一个半球的能力及方式去工作。

例如，当你在读一个故事的时候，大脑右半球正专注于故事的感情方面（幽默、悲伤），想象其视觉情景，保持故事整体结构的印象，并且欣赏其艺术写作手法（例如隐喻的运用）。当这些活动发生的时候，大脑左半球正在理解书面语言，从词句之间的复杂关系中寻找意义，并且将这些词转换成语音，使它们可以作为语言来理解。你能阅读、理解并欣赏一个故事，是因为大脑是作为一个单一、整合的结构运行的（Levy，1985）。

利维认为，事实上，没有一项人类活动是只用大脑的一个半球来完成的：“那种流行的神话只是一种解释和愿望，并不是科学家的客观发现。正常人不会只有半个脑，也不会有两个脑，有的只是一个分化的大脑，大脑的每个半球都具有特定的功能。”（Levy，1985，P.44）

近期应用

对利维见解的引用再次说明了斯佩里和加扎尼加割裂脑研究的持续影响。近期一篇评论医学和心理学文献的文章指出，有许多领域的大量文章不仅参考了罗杰·斯佩里早期的研究及方法，而且也有很多参考了加扎尼加及其合作者的最新研究成果。例如，1998年，法国的一项研究（Hommet & Billard，1998）对斯佩里和加扎尼加的研究基础提出质疑，认为切断胼胝体难以把大脑分成两半。法国的这一研究发现，对于一种天生就没有胼胝体的孩子（一种罕见的脑畸形），信息仍可以在他们大脑的两半球之间传递。研究者推断，这些孩子脑内一定存在着除胼

胝体以外的其他形式的重要连接方式。在割裂脑患者中是否存在这样的皮层下联系现在还不清楚。

最近的研究提示，在将加扎尼加的研究成果应用于教育时需要特别注意（Alferink & Farmer-Dougan，2010）。一种广泛流行的观点是大脑两半球控制着明显不同的认知功能。这一结论来自于因病而进行外科手术、切断胼胝体的少数病人。我们不能没有限制地将这些结论用于大脑正常的人身上。基于大脑半球负责特定任务加工观点的教育模式将是非常危险的。一些研究者指出，上述研究中的病人即便在手术前，其大脑功能也是异常的，因此，在针对正常人的教育上，我们应该避免只关注大脑的一个半球。

尽管如此，研究者仍在继续探究关于大脑两半球彼此独立的不同功能与影响的问题。一个研究（Morton，2003）揭示了优势半球是如何决定个体特定的兴趣和职业的。Morton的研究在这方面有两个发现：第一，通过使用最佳利手测验（The Best Hand Test）测量脑半球偏侧化（即一个人是左脑取向还是右脑取向），Morton发现在上大学一年级基础课的400名学生中，56%的人是左脑取向。将同样的方法用于另外180名学习特殊高阶课程的学生，发现左脑取向的学生在38%到65%之间，这一差异意味着大脑半球可能与大学生学业成就和兴趣有关；第二，Morton使用相同的方法探讨在大学中，大脑半球取向是如何决定不同专业岗位的。研究结果发现，大脑半球偏侧化能够显著地预测个人的职业选择。在Morton的研究中，83%的生物化学家是左脑取向者，天文学家中有29%的人是左脑取向者（P.319）。从这项研究中，你可以发现其与斯佩里和加扎尼加的工作有某种联系，生物学家和化学家更多地依赖于其语言能力，而天文学家必须在空间关系方面具备更好的能力。

结论

一些研究者进一步发展了上述观点，并将其应用于一些心理紊乱问题上，如分裂性多重人格障碍（见Schiffer，1996）。其背后的思想是，对于一些具有完整的、“非割裂”大脑的人，他们的大脑右半球可能会比左半球的功能更独立，甚至可以在一段时间内控制一个人的意识。多重人格障碍有没有可能是隐藏于大脑右半球的人格的外在表现呢？这是一个值得用你大脑的两个半球思考的问题……

Alferink, L., & Farmer-Dougan, V.(2010).Brain-(not) based education: Dangers of misunderstanding and misapplication of neuroscience research.Exceptionality，18, 42-52.

Gazzaniga, M.S.(1985).The social brain.New York:Basic Books.

Gazzaniga, M.S., & Ledoux, J.E.(1978).The integrated mind.New

York: Plenum Press.

Hommet, C., & Billard, C.(1998).Corpus callosum syndrome in children.Neurochirurgie, 44(1)，110-112.

Levy, J.(1985, May).Right brain，left brain: Fact and fiction.Psychology Today, 42-44.

Morton, B.E.(2003).Line bisection-based hemisphericity estimates of university students and professionals:Evidence of sorting during higher education and career selection.Brain and Cognition, 52(3), 319-325

Puente, A.F.(1995).Roger Wolcott Sperry (1913-1994).American Psychologist, 50(11).940-941.

Schiffer, F.(1996).Cognitive ability of the right-hemisphere: Possible contributions to Psychological function.Harvard Review of Psychiatry, 4(3)，126-138.

Sperry, R.W.(1968).Hemisphere disconnection and unity in conscious awareness.American Psychologist, 23, 723-733.

注释

[1]玛丽·雪莱创作的科幻小说《科学怪人》中的主人公，他通过人工手段制造出生命，而非真正的科学。——译者注

研究2 丰富的经历是否等于更大的大脑

Rosenzweig, M.R., Bennett, E.L., & Diamond, M.C.(1972).Brain changes in response to experience.Scientific American, 226 (2), 22-29.

如今，如果你进入一个典型美国中产阶级家庭的婴儿房间，你可能会看到一张婴儿床，在婴儿伸手可及的上方悬挂着许多小动物玩具和各种各样的彩色玩具。其中一些玩具会发光、会活动、会演奏音乐，或者兼有上述三种功能。设想一下，人们为什么给婴儿准备这么多可看可玩的东西呢？除了婴儿喜欢这些东西并作出积极反应外，绝大部分父母都相信，无论这个观点得到公认与否，孩子们需要一个令他们兴奋的环境，以促进智力和大脑最大程度的发展。

关于某种特定的经历是否会引起大脑形态变化的问题，几个世纪以来一直是哲学家和科学家在猜测和研究的话题。1785年，意大利解剖学家玛拉卡尼（Malacarne）研究了同一胎产下的几对小狗和同一窝蛋孵出的几只小鸟。他有目的地长期训练每一对中的一只，而另外一只会得到同样良好的照料，但并不接受训练。然后，通过对动物的尸体进行解剖，他发现受过训练的动物的大脑表现得更为复杂，带有更多的褶皱和沟回。然而，这一研究不知因为什么原因没能继续下去。在19世纪后期，人们试图把一个人的学习量和他脑袋的周长联系起来。虽然一些早期的研究成果支持这种相关，但后来的研究成果认为这并不是一种测量大脑发展的有效尺度。

直到20世纪60年代，新技术的发展使科学家们具备更精确地检测大脑变化的能力，他们运用高倍显微技术，并对大脑内各种酶和神经递质水平进行评估。在加利福尼亚大学伯克利分校，马克·罗兹维格（Mark Rosenzweig）及其同事爱德华·本奈特（Edward Bennett）以及玛丽安·戴蒙德（Marian Diamond）采用这些技术，历时10余年，进行了由16次实验组成的系列研究，力图揭示经验对大脑的影响。本章的这篇文章将介绍他们的研究发现。由于显而易见的原因，在他们的研究中并没有用人作被试，而是像很多经典心理学实验一样，用老鼠作被试。

理论假设

由于心理学家最终的兴趣在于人而不是老鼠，因而就必须指出这种不用人作被试的研究的合理性。在这些研究中，为什么选择老鼠作被试就成了研究理论基础的一部分。作者解释说，由于多种原因，使用啮齿类动物比使用高级的哺乳类动物（如食肉类或灵长类动物）更方便。这

项研究的重点是脑部，老鼠的脑部是平滑的，并不像更高等的动物那样曲折而复杂。因此，对其大脑的检测和测量就更容易。此外，老鼠体型较小并且也不昂贵，在实验室的研究中，这是一个很重要的考虑因素（通常实验室经费并不充足而且缺乏空间）。老鼠一胎多子，这就允许研究者将同一窝中的老鼠分配到不同的实验条件下。作者最后指出，研究者培养了多种种系的老鼠，以便在需要的时候把遗传因素考虑在内。

在罗兹维格的研究中，隐含着一种想法，即把饲养在单调或贫乏环境中的动物与饲养在丰富环境中的动物进行对比，二者在大脑发育和化学物质等方面将表现出明显的不同。在这篇实验报告所涉及的每次实验中，均采用了12组老鼠，每一组均由取自同一胎的三只雄鼠组成。

方法

三只雄鼠都是从一胎所生的老鼠中选择的，它们被随机分配到三种不同的实验条件中。一只老鼠仍旧与其他同伴待在实验室笼子里，另一只被分派到罗兹维格称为“丰富环境”的笼子里，第三只被分派到“贫乏环境”的笼子里。记住，在16次实验中，每次都有12只老鼠被安排在每一种实验条件中。

三种不同环境（如图1—4）描述如下。

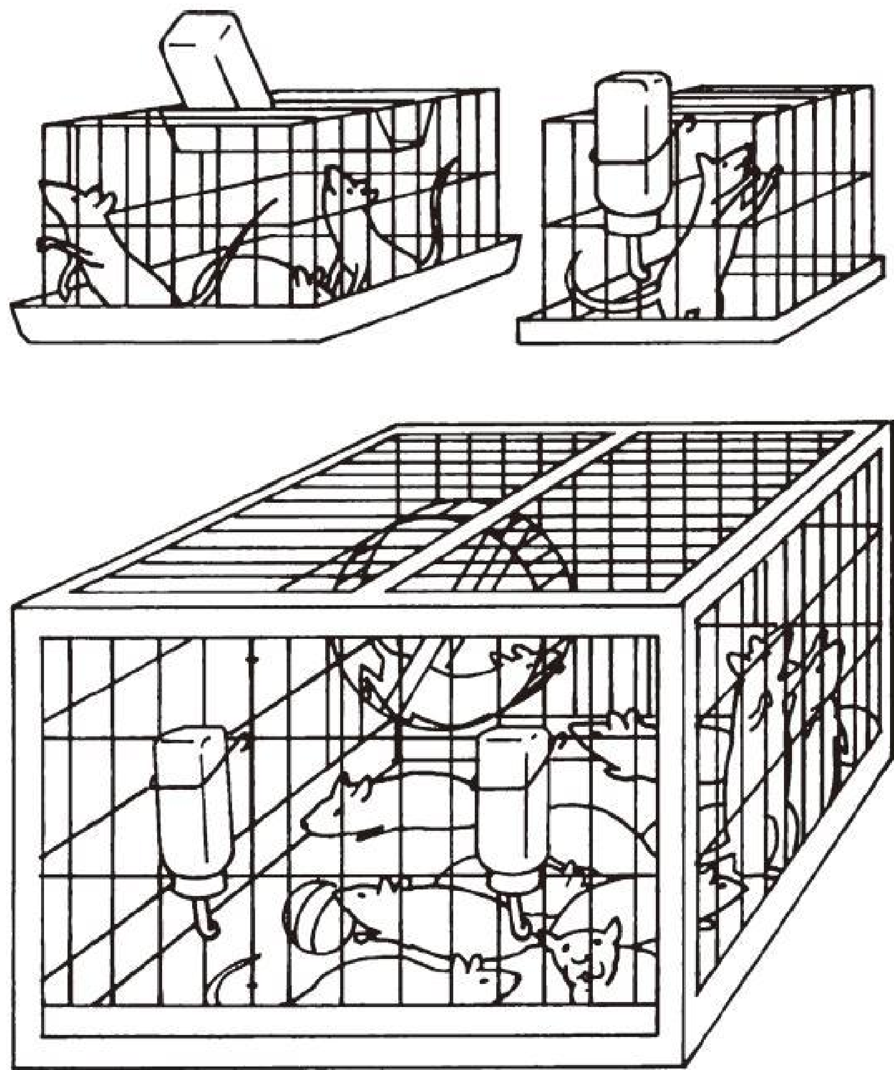


图1—4 罗兹维格等人实验中三个笼子的环境

1.标准的实验室笼子中，有几只老鼠生活在足够大的空间里，笼子里总有适量的水和食物。

2.贫乏的环境是一个略微小一些的笼子，老鼠被放置在单独隔离的空间里，笼子里总有适量的水和食物。

3.丰富的环境几乎是一个老鼠的迪士尼乐园（并没有冒犯米老鼠的意思），6~8只老鼠生活在一个“带有各种可供玩耍的物品的大笼子里。每天从25种玩具中选取一种新的组合放在笼子里”（第22页）。

实验人员让老鼠在这些不同环境里生活4周到10周不等。经过这样

不同阶段的实验处理之后，实验人员将人道地剥夺这些实验老鼠的生命，对它们进行解剖以确定脑部发育是否不同。为了避免实验者偏见的影响，解剖按照编号的随机顺序进行，这就可以避免尸检人员知道老鼠是在哪种环境下成长的。研究者关注的是生活在丰富环境下与生活在贫乏环境下的老鼠的大脑所出现的不同。

解剖老鼠的大脑后，对各个部分进行测量、称重和分析，以确定细胞生长量与神经递质活动的水平。在对后者的测量中，有一种叫作“乙酰胆碱”的脑酶引起了研究者们特别的兴趣。这种化学物质十分重要，因为它能使脑细胞中神经冲动传递得更快、更高效。

罗兹维格和他的同事是否发现了老鼠的大脑因其生活环境的丰富或贫乏而有所不同呢？以下就是他们的研究结果。

结果

结果证实，生活在丰富环境中的老鼠，其大脑在很多方面都有别于贫乏环境中的老鼠。在丰富环境中生活的老鼠的大脑皮层（即大脑对经验作出反应的部分，它负责行动、记忆、学习和所有感觉的输入，如视觉、听觉、触觉、味觉、嗅觉）更重且更厚，并且这种差别很显著。前面提到神经系统中存在的“乙酰胆碱”酶，在身处丰富环境的老鼠的大脑组织中更具活性。

两组老鼠的脑细胞（又称为神经元）在数量上并没有显著差别，但丰富的环境使老鼠的大脑神经元变得更大。与此相关，研究还发现，丰富环境中的老鼠的RNA和DNA（两种对神经元生长起最重要作用的化学成分）的比率也相对更高。这意味着在丰富环境里长大的老鼠其大脑中有更高水平的化学活动。

罗兹维格和他的同事解释说：“虽然由环境引起的大脑变化并不很大，但我们确信这种变化是千真万确的。在重复实验的时候，上述结果仍能出现……我们发现，经验对大脑最一致的影响表现在大脑皮层与大脑的其余部分（即皮层下部）的重量之比上。具体表现为，经验使大脑皮层迅速地增重，但大脑其他部分变化很小”（P.25）。这种对大脑皮层与大脑其余部分比率的测量是对大脑变化最精确的测量。这是因为每只动物的脑重量会随动物体重的变化而变化，运用这个比率，可以消除个体差异。图1—5用图表说明了16次实验的结果。正如你看到的，只有一次实验结果的差异在统计上不显著。

最后，是有关两组老鼠大脑的神经突触的发现。神经突触是指两个神经元相汇之处。大部分大脑活动发生在神经突触上，在这里，神经冲动有可能通过一个又一个神经元继续传递下去，也有可能被抑制或终止。在高倍电子显微镜下，能发现在丰富环境中长大的老鼠其大脑中的

神经突触比在贫乏环境下长大的老鼠的神经突触大50%。

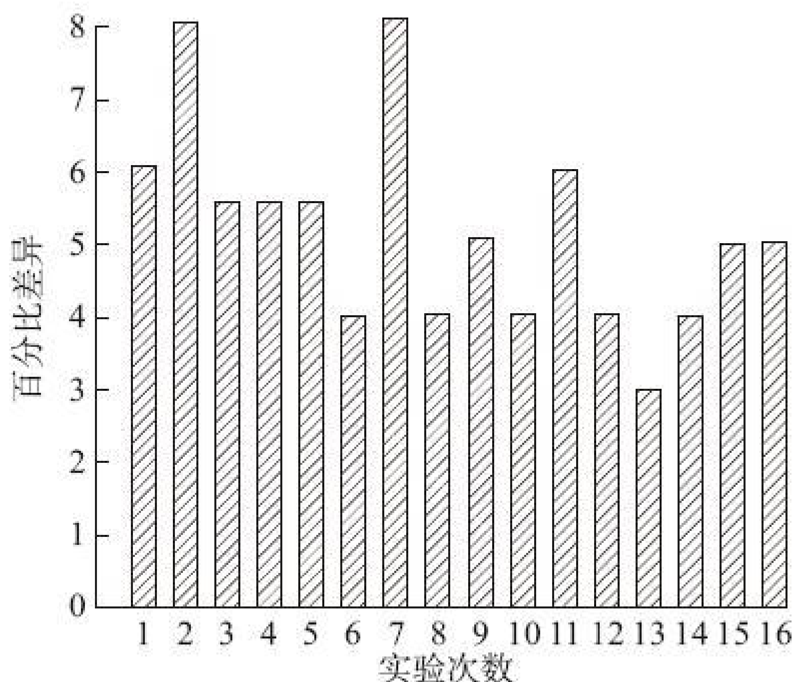


图1—5 丰富环境与贫乏环境相比老鼠大脑皮层增加的比率

注：实验2至16有统计学上的显著差异。

资料来源：引自Rosenzweig, Bennett, & Diamond, P.26。

讨论及批评

经过近10年的研究，罗兹维格、贝奈特和戴蒙德满怀信心地指出："毫无疑问，大脑构造及其化学成分的很多方面可以被经验改变。"（P.27）。然而，他们也坦承，当他们第一次公布研究结果时，许多科学家对此表示怀疑，因为像这样的结果在过去的研究中从未被清楚地证实过。有些批评意见认为，也许并不是丰富的生活环境使大脑产生了变化，而是由于老鼠接受了不同的实验处理，如纯粹的触摸或压力情境。

这种批评是有根据的，在丰富环境的实验条件下，老鼠每天被触摸两次，即在换玩具时得将它们移开。而在贫乏环境的实验条件下，老鼠没有被触摸。因此，可能是触摸导致了这一结果，而不是丰富环境的实验条件所致。为了消除这一潜在的混淆因素，研究者每天抚摸一组老鼠，而不抚摸与它们同胎的另一组老鼠（这些老鼠都处在同一环境

中）。罗兹维格和他的同事并没有发现这两组老鼠的大脑有什么不同。此外，在他们后来的研究中发现，同样抚摸在丰富环境与贫乏环境条件下的老鼠，得到的结果也是一样的。

至于压力情境，批评者认为，在贫乏环境中生活的老鼠因被隔离而产生压力，这是导致它们大脑欠发达的原因。罗兹维格等人引证了另一项研究，此项研究让老鼠暴露于日常的压力情境之中（旋转笼子或给予轻微电击），并没有发现仅因压力而使大脑改变的证据。

在实验室中进行的任何研究都存在人为性问题。罗兹维格和他的同事很想知道在自然的生长环境中，各种水平的刺激是如何影响动物的大脑发育的。他们指出，实验室中的老鼠常在人工环境中繁殖，且已经繁衍了100代，它们和野生鼠几乎是没有相似的遗传基因。为了探索这种有趣的可能性，他们开始研究野生老鼠。他们把抓到的这种野生鼠随机地放在户外自然环境中或是实验室的丰富环境笼子里。4周后，发现户外老鼠的大脑比实验室老鼠的大脑发展得更好。“这就表明，实验室中的丰富环境与自然环境相比，仍是相当贫乏的”（第27页）。

最后，涉及动物被试最主要的批评意见是它与人的关系问题。毫无疑问，这类研究永远不能用人来做被试，但是研究者仍有责任探讨此问题，并且这些科学家也这样做了。

作者解释说，很难把在一组老鼠身上的研究结果泛化到另一组老鼠身上，而要把用老鼠做研究的结果泛化到猴子或人的身上则更难。虽然他们宣布在几类啮齿类动物身上取得了相似的结果，但他们也承认，在得出经验对人脑产生影响的假设前，仍需要更多研究。然而，他们提出以动物为被试的这类研究的价值在于“允许我们对概念和技术进行检验，其中一部分可能对今后以人为被试的研究有所帮助”（第27页）。

作者在这篇文章中还提到这项研究的几个潜在好处。它可以用在对记忆的研究中。由于经验而使大脑产生的改变，可以使人更好地理解记忆是怎样被保存在大脑中的，从而导致一些新技术的产生以提高记忆并阻止因年龄增长而导致记忆的衰退。另一方面，这些研究也有助于解释营养不良与智力发展之间的关系。作者的观点是，营养不良可以使一个人对环境刺激的反应迟钝，长期持续的营养不良可能限制大脑的发育。作者强调，一些同时进行的研究证明，营养不良对大脑发育的影响也会因环境的丰富而减轻，或因环境的恶劣而加重。

相关研究和近期应用

这项由罗兹维格、本奈特和戴蒙德所做的工作，对这一领域研究的持续发展起到了催化剂般的作用。自他们论文发表的25年以来，这几位科学家和其他科学家一道继续致力于巩固、改进、扩充他们的研究成

果。

例如，他们发现丰富环境的生活经历有助于提高学习本身，即使是在贫乏环境中长大的成年动物，当它被放在一个丰富环境中的时候，其大脑发育也能获得提高（见本奈特，1976，一个完整的回顾）。

现有的一些证据表明，经验确实改变了人类大脑的发育。通过对自然死亡者的尸体解剖，研究者发现当一个人具有更多的技术和能力时，他的大脑确实变得更复杂也更重。在对那些没有某种特定经历的人的大脑进行解剖时，他们发现了另外一些结果。例如，与视觉正常的人相比，盲人大脑的视觉皮层部分明显发育不良，沟回较少，皮层较薄。

玛丽安·戴蒙德（Marian Diamond）是原稿的作者之一，她在人类智力的毕生发展进程方面运用了这一研究成果。她说："在生活中，我想我们应该对大脑老化持一个更乐观的态度……主要因素是刺激，神经细胞因刺激而存在，所以我认为好奇心是一个关键因素。如果一个人在一生中始终充满好奇心，那这将刺激他的神经组织和大脑皮层……我寻访在88岁以后还极有活力的人。我发现那些经常用脑的人，大脑不易老化。事情就是这样简单。"（Hopson，1984，P.70）

两项新近研究从不同的实际应用角度阐述了罗兹维格、戴蒙德和本奈特关于环境影响大脑发育的观点。怀斯和贝林格（Weiss & Bellinger，2006）将研究扩展到探讨有害环境对人类大脑早期发展的影响，他们认为有害环境不仅应包括各种有害的化学物质，还应考虑个体生活情境中的所有因素，如遗传倾向和生活环境的丰富或贫乏。研究者提出，对人类而言，接触有害物质造成的影响与个体成长环境的丰富或贫乏直接相关。换言之，在贫困中长大的儿童，不仅其发展环境相对贫乏，而且其暴露于有害物质中的风险更高。环境因素甚至会作用于有害物质，从而对大脑发展造成影响。怀斯和贝林格称，对环境中的有害物质的研究已呈现出一种趋势，即集中于研究有害物质本身，而淡化相关的情境变量。作者写道：

我们认为，生命早期暴露于具有神经毒性化学物质中所造成的后果，取决于儿童的社会环境，包括其出生前的主导环境……我们认为，要对有害物质在神经行为层面的潜在后果作真正评估，离不开个体活动、导致其独特性和持续造成个体缺陷的生态环境（如环境丰富）（P.1497）。

另一篇文章也引用了罗兹维格在1972年的研究，对近来一些试图用过分简单化的丰富环境策略来促进儿童大脑发育的做法提出了批评（Jones & Zigler，2002）。正如你料想的那样，当公众了解到罗兹维格等人的研究后，就有可能掀起一些看似极有吸引力的热潮，但它们却极少建立在科学事实之上。20世纪90年代出现的"莫扎特效应"（Mozart Effect）就属其中一例。这一股狂热开始于一些初步研究所发现的结果：儿童听莫扎特（而不是其他古典作曲家）的音乐，学习将变得更

好。这一思想在互联网上发展到了极致，“莫扎特效应”可使儿童和成人从中获益，并宣称特定的音乐能够促进全身健康、提高记忆力、治疗注意缺陷、减少抑郁和加速身体创伤的愈合。

结论

琼斯和奇格勒（Jones & Ziger，2002）认为这些流行的应用对研究的推广不仅无效而且危险。他们指出：“将脑研究的结果作为一种‘权宜之计’，用来解决复杂系统的问题，这是不恰当的。”（P.355）他们进一步提出，当科学的脑与学习的研究得到谨慎的、正确的应用时，它才能够“对早期认知和社会性发展起到深入的、高质量的、多领域的干预作用”（P.355）。

Bennett, E.L.(1976).Cerebral effects of differential experience and training.In M.R .Rosenzweig & E.L.Bennett (Eds.), Neural mechanisms of learning and memory.Cambridge, MA: MIT Press.

Hopson, J.(1984).A love affair with the brain: A PT conversation with Marian Diamond.Psychology Today，11, 62-75.

Jones, S.& Zigler, E.(2002).The Mozart Effect: Not learning from history.Journal of Applied Developmental Psychology, 23, 355-373.

Weiss, B., & Bellinger, D.C.(2006).Social ecology of children’s vulnerability to environmental pollutants (Commentary).Environmental Health Perspectives，114，1479-1485.

研究3 人的本性是“天生的”吗

Bouchard, T., Lykken, D., McGue, M., Segal, N., & Tellegen, A.(1990).Sources of human psychological differences: The Minnesota study of twins reared apart.Science, 250, 223-229.

近来，在如何广义地看待人类行为这一问题上，心理学家的意见正发生着根本性的变化，本项研究正是这一变化的代表。倘若你想体会一下这种变化，可以对以下这个问题稍作思考：“你是怎样的一个人？”只需花一点儿时间思考一下自己的个性特征，即你的“人格特质”：你是易紧张激动还是平静放松？内向腼腆还是开朗大方？富于冒险精神还是追求舒适安逸？合群还是孤僻？对未来是乐观还是悲观？请思考一下这些问题以及其他你认为与之有关的问题。慢慢儿来……完成了吗？现在，请回答下面这个对本文而言非常重要的问题：“你为什么是这样的一个人？”换言之，是什么因素使你变成了现在的你？

一般而言，你会和大多数人一样，将此归结于父母对子女的教育以及他们灌输给子女的价值观、人生观和权衡轻重的原则等。你也许会认为这其中有兄弟姐妹、祖父母、叔叔阿姨以及在你的品格塑造方面起关键作用的同伴、老师和其他良师益友的影响；有些人还会想到生命中的一些重大转折，如生病，失去自己所爱的人，对大学、专业甚至是对特殊生命历程的选择，都有可能对一个人人格的塑造产生极大影响。以上这些只是众多影响因素中的一部分。然而，所有这些影响都具有一个共同的特征：那就是它们都是环境因素。几乎没有人会对“你为什么是这样的一个人”这个问题报以“我生来就是这样”、“这都是我的基因使然”的回答。

所有人都承认身高、发色、体型及眼睛的颜色等体态特征具有遗传性。且越来越多的人开始认识到癌症、心脏病和高血压等许多疾病的发病倾向也有很明显的遗传成分。但几乎没有人会想到在人的心理品质中，基因也起着很重要的作用。说到这些，你一定会感到很奇怪，但在我们的“环境偏向”事实上的确有原因可循。

首先，在20世纪后半叶，心理学领域一直被一种叫行为主义的人性理论所统治。行为主义理论的基本主张是人的行为被环境因素所左右，其中包括引起人的行为的刺激和某种选择反应的结果。严格的行为主义者相信，人的内部心理活动不仅不能用科学的手段加以研究，而且这种研究对完整解释人类的行为无关而多余。行为主义的理论是否为大众文化所理解或接受其实并不重要，重要的是，其对如今根深蒂固的大众信仰的现实影响，大家普遍认为，经验是人性塑造的基本因素或唯一因素。

其次，遗传和生理因素对人类行为的影响不能直接观察到，这便是人们普遍接受用环境因素解释人类行为的又一个原因。人们常会说：“我成了一名作家，这是因为在我七年级时，我的写作老师曾给予我深深的启发和鼓励。”你记得这些影响的存在，它们是你亲眼所见的，是你过去生活中的一部分，还是你目前能够意识到的经验。我们很难意识到生理因素的影响，很少有人会说：“我成了一名作家，是因为在我的DNA中包含了一种基因，这种基因在我的身上得以表现，它使我先天就具备了妙笔生花的倾向。”基因的影响看不见，摸不着，甚至我们根本不知道它位于身体的哪个部位！

最后，很多人不愿接受基因在其人格塑造中起决定作用的观点，而宁愿相信真正起作用的是他们在生活中所作出的各种选择。因为前者带有宿命论的味道，而且无法体现人的“自由意志”。对那些可能会限制个体就生活中的一些结果作出决定的意志力的理论，很多人都很反感。因此，他们便趋向于回避和反对行为和人格的遗传学解释。事实上，完整的人格是通过遗传和经验的相互作用而塑造起来的，我们所要问的唯一的一个问题是：此二者中，哪个占了主导地位？或用媒体的惯用表达方式来说，便是：“人性是先天的还是后天的？”

本节我们将要探讨的是由托马斯·鲍查德（Thomas Bouchard）、大卫·莱肯（David Lykken）及其助手们在位于明尼阿波利斯市的明尼苏达大学所完成的论文。文章所涉及的是一项开始于1979年的研究，该研究意在检验基因在决定个人心理品质中所起作用的大小。人们试图用一种科学的方法将人的行为和人格中的遗传影响（先天）和环境影响（后天）加以分离，这便促使了该项研究的出现。但是，假定你们每个人都不是被收养的，且生长于基因赠予者（即父母）所营造的环境里，你们想想，此时要把遗传和环境因素分离开来就不那么简单了。举个例子，你可能和你父亲一样具有幽默感（绝无冒犯之意），这种幽默感是你从他那儿学到的（后天的）呢，还是你继承了他那载有“幽默感”的基因（先天的）？我们似乎无法找到一种系统的方法来把这两种因素分开，对吗？

然而，鲍查德和莱肯会回答“错！”他们发现有一种方法可以确定某种心理特征到底是由遗传因素还是由环境因素决定的，他们对这种方法颇有信心。

理论假设

这确实不难，你所要做的就是找到两个具有相同基因的人，从出生起就把他们分开，在两种截然不同的环境中把他们抚养成人。然后，你就可以假定，到他们成年时，他们的行为和人格的相同之处便是遗传因素所致。但研究者到底怎样才能找到两个一模一样的人呢（别说是“克隆”；我们还没到那一步）？即使我们能找到，要强迫他们进入完全相反

的生存环境也很不道德，不是吗？正如你已经猜到的那样，研究者没必要这么做。社会已经为他们做好了这一切：同卵双胞胎实际上就具有完全一致的遗传结构。之所以称其为同卵双胞胎是因为他们始于同一个受精的卵子，简称受精卵，然后才分裂成为两个相同的胚胎。异卵双胞胎是两个不同的精子细胞使两个不同的卵子受精后的结果，所以把他们称为异卵双胞胎。异卵双胞胎和其他非双胞胎的兄弟姐妹一样，仅具有遗传的相似性。不幸的是，双胞胎弃婴被不同家庭所收养的情况时有发生。收养机构也会尽力把兄弟姐妹放在一起，特别是双胞胎，但更重要的是，要为他们找到一个好的家庭，尽管这有时会意味着分离。所以，随着时间的推移，数以千计的同卵、异卵双胞胎为不同的家庭所收养，他们在不同的、有时甚至是截然不同的环境下长大，而且他们通常都不知道自己还有个双胞胎的兄弟姐妹。

鲍查德和莱肯从1983年便开始鉴定、寻找这类双胞胎并将他们集中起来，他们最终找到了56对分开养育的同卵双胞胎（MZA）。这56对双胞胎来自美国等8个国家，他们同意参加为期一周的集中心理测验和生理测量。在进行进一步研究分析后，鲍查德等人于1990年完成了这篇研究报告（这项研究是在“双子城”之一的明尼阿波利斯市进行的，它无法不引起了人们的关注）。研究者将这些双胞胎与那些共同成长的同卵双胞胎（MZT）进行比较，得出了惊人的发现，在整个生物与行为科学领域引起了巨大的反响。

方法

被试

这项研究所面临的第一个挑战就是要寻找那些早年分离、成长环境不同、成年后才相聚的同卵双胞胎。研究者要进行这项研究的消息以口头传播方式流传开来以后，找到了许多被试。双胞胎本人、朋友或家庭成员与明尼苏达双胞胎收养和研究中心（MICTAR）取得联系，该机构里从事各种社会公益事业的专业人员也在其中协助进行联系工作。有时，也会出现双胞胎之一与中心取得联系并寻求帮助的情况，他们希望找到自己的兄弟姐妹。所有的双胞胎在参加研究之前，均经过检测，以确保他们确实是同卵双胞胎。

程序

研究者想要在双胞胎来访的一周内获得足够多的资料。每一名被试完成将近50小时的测试，测试内容几乎涵盖你所能想到的每个维度。他们完成了4种人格特质量表、3种能力倾向和职业兴趣问卷以及两项智力测验。另外，被试还要填写一张家用物品清单（例如家用电器、望远镜、艺术真品和辞海等），以评估其家庭背景的相似性；一张家庭环境

量表以测量他们对养父母教养方式的感受。他们还要进行个人生活史、精神病学以及性生活史等方面的访谈。每名被试的所有项目全部分开独立完成，以避免一对双胞胎间存在不经意的相互影响。

正如你所想到的，数小时的测验产生了数量巨大的信息资料。下面我们将讨论这其中最重要、最令人惊讶的结果。

结果

表1—1显示了分开养育的同卵双胞胎（MZA）在某些特征上的相似性，也包含了共同养育的同卵双胞胎（MZT）在该方面的测量结果。相似程度在该表中用相关系数或相关值“r”来表示。相关系数越大，其相似程度越高。在此，有这样一个逻辑假设：若个体的差异是由环境引起的，则在相同环境下成长起来并共同养育的同卵双胞胎与分开养育的同卵双胞胎相比，前者的个体特征应更相似。正如你所看到的，研究者所发现的并非如此。

将MZA双胞胎间每种特征的相关系数与MZT双胞胎间的相关系数相除，所得数值列在表1—1的最后一列，这列数值表示两类双胞胎在每种特征相似性上的差异。如果两个相关系数相同，则相除以后的结果是1.00；如果它们完全不同，则相除以后的结果会接近0.00。仔细观察表1—1第4列中的数据，就会发现MZA双胞胎和MZT双胞胎在每种特征上的相关系数惊人地相似，即其比值大多接近于1.00，没有低于0.70的。

分开养育的同卵双胞胎（MZA）与共同养育的

表1—1 同卵双胞胎（MZT）在某些特征上的相关系数（r）的比较*

特征	r（MZA）	r（MZT）	相似性 r（MZA）÷r（MZT）**
生理	—	—	—
脑电波活动	0.80	0.81	0.987
血压	0.64	0.70	0.914
心率	0.49	0.54	0.907
智力	—	—	—
韦氏成人智力量表	0.69	0.88	0.784
瑞文智力测验	0.78	0.76	1.030
人格	[illegible]	[illegible]	[illegible]
多维人格问卷（MPQ）	0.50	0.49	1.020
加利福尼亚人格问卷	0.48	0.49	0.979
心理兴趣	—	—	—
史特朗—康久尔兴趣问卷	0.39	0.48	0.813
明尼苏达职业兴趣量表	0.40	0.49	0.816
社会态度	—	—	—
宗教信仰	0.49	0.51	0.961
无宗教信仰社会态度	0.34	0.28	1.210

*（选自Table 4，p.226）。

**1.00意味着MZA双胞胎之间的相关与MZT双胞胎之间相关极其相似。

讨论及研究发现的意义

这些结果表明，对于相当数量的人类特征而言，大多数差异似乎是由遗传因素（或“基因”）引起的。表1—1中的数据从两个重要方面证明了这一结果。其一是，具有完全相同的遗传特质的人（同卵双胞胎），即便分开抚养，且生活条件大相径庭，在他们长大成人以后，不仅在外表上极为相似，而且其基本心理和人格也惊人地一致。在本研究中，支持基因占主导地位的第二个论据是，在相同条件下养育的同卵双胞胎，环境对他们的影响似乎很小。鲍查德和莱肯将他们的发现表述如下：

到目前为止，在调查过的每一种行为特征——从反应时间到宗教信仰，个体差异中的重要部分都与遗传有关。这一事实今后不应再成为争论的焦点；现在是该考虑它的意义的时候了。

当然，对于鲍查德和莱肯的这一观点，仍然有人会提出疑议。我们将在下一部分对不同的观点进行讨论。然而，探讨本项研究及其研究者所作的类似研究的意义是绝对有必要的。本文所报告的遗传学发现以何种方式改变了心理学家，甚至改变了我们每个人对人性的看法？正如前面提到的，心理学界和西方文化被环境决定论的思想统治了近50年。过去，我们对于教育、犯罪和刑罚、心理治疗、技能和能力、兴趣、职业目标、社会行为等许多问题持环境决定论的解释，认为是人的经验而非

基因塑造了其人格。很少有人会以“那个人生来就是这样”的想法来理解某人的行为。我们确信，人的行为模式是习得的，因为我们看到，父母的不同养育方式造成了孩子人格上的差异，积极的生活体验可以战胜消极的生活体验，不健康的无效行为可以得到消除。而“人格是先天的”这样的观点却诱使我们提出一系列以“何必”开头的问题：“何必努力成为好家长？”“何必去帮助那些穷困潦倒的人？”“何必努力提供高质量的教育？”等等。那么，鲍查德和莱肯便是第一个站出来，反对对其研究发现作这种解释的人。在这篇文章中，他们对自己那些具有煽动性的结论提出了以下三种解释。

1.显然，智力主要是由遗传因素决定的（智力差异中的70%都可以归因于遗传的影响）。然而，正如作者所明确阐述的：

> 这些发现并不意味着像IQ这样的特质是不能提高的……一项覆盖了14个国家的调查表明，近年来IQ测验的平均成绩已有所提高。因此，目前的发现并没有规定和限制IQ成绩在理想环境中的提高（P.227）。

从根本上讲，他的意思是，虽然IQ差异中的70%要归因于先天基因的不同，但仍有30%可归因于环境的影响。这些影响包括许多我们众所周知的因素，例如教育、家庭环境、毒品和社会经济地位等。

2.鲍查德和莱肯研究的基本假设是，人的特性是由遗传和环境的综合影响决定的。所以，当环境因素影响较小时，其差异更多地来自遗传，反之亦然：一方面，对某些特性而言，如果环境因素对其影响较大，则遗传的影响就较小，例如，在美国，大多数儿童都有机会学习骑自行车，这就意味着对所有儿童而言，环境的作用几乎是很相近的，所以在骑车能力方面的差异主要受遗传因素的影响；另一方面，在美国，人们在食物偏好上的不同，似乎受环境因素的影响较多，因为在童年和整个人生中，各人所接触的食物或口味有很大不同，因此，几乎没有给遗传因素留有发挥其作用的余地。在此，研究者提出了一个很有意思的观点，即他们认为人格更像是骑自行车，而不像是对食物的偏好。

实际上，作者的意思是，在孩子长大成人的过程中，家庭环境对他的影响与他所继承的基因相比，其作用是很小的。而大多数家长不愿意听信这些，这是可以理解的。他们努力扮演好父母的角色，并使孩子快乐成长，让他们成为一名好公民。而唯一能从这些研究结果中得到安慰的是那些在抚养子女方面已经黔驴技穷的家长，这样的研究结果能使他们少一些自责！然而，鲍查德和莱肯很快又指出，基因不是决定命运的必然因素，那些极富奉献精神的父母仍能从正面影响他们的子女，即使他们在总体变异中所占的百分比很少。

3.鲍查德和莱肯提到的最有趣的一点是，并非环境影响人的特性。恰恰相反，是人的特性影响着环境。也就是说，实际上是人的遗传倾向塑造着周围环境！有这样一个例子可以说明这种理论：有些人比其他人更具有丰富的情感，人们常认为这是前者的父母对孩子比后者的父母更亲切慈爱的结果。换言之，情感丰富的孩子来自富于情感的环境。每当

对这种假设进行研究时，总能发现这种说法是正确的。的确，情感丰富的孩子从他们父母那儿得到了更多的爱。然而，鲍查德和莱肯认为，事实上，“感情丰富性”方面的差异是由遗传决定的，所以有些孩子一出生就比别的孩子情感更丰富。与别的孩子相比，他们这种与生俱来的行为倾向，使其能够对父母的爱作出更主动的反应，这在某种意义上就强化了其父母的行为。正是这种反应引起了父母充满爱心的行为。研究者强调，即便遗传不是决定人的绝大部分特性的因素，但它也在许多人类特性中发挥着重要作用。他们这样认为：

> 导致绝大多数心理差异的直接原因可能包含着通过经验而习得的东西，这一点也是极端的环境论者所坚信的。然而，有效的经验在很大程度上是人们自我选择的结果，而这种选择又为遗传的恒定压力所引导（P.228）。

批评及相关研究

正如你所想象的，人们以鲍查德和莱肯所得的双胞胎数据资料为基础，完成了大量的相关研究。这些研究结果仍表明，基因对许多人格特征和行为的影响是很大的。通过双胞胎研究，我们需要对一些过去在很大程度上或完全归因于环境的特性进行重新评价，正如双胞胎研究所揭示的，遗传因素不仅是变化的重要因素，而且其作用比我们以前所预料的还要显著。

例如，明尼苏达大学的研究小组发现，基因不仅在很大程度上决定着人们对职业的选择，甚至当各种职业所要求的生理条件保持恒定时，在人们的工作满意度和职业道德方面大约有30%的变化缘于遗传因素（Arvey et al.，1989；Arvery et al.，1994）。另有一项研究对同卵双胞胎和异卵双胞胎进行比较，这些双胞胎中，既有共同抚养的，也有分开抚养的，研究者把研究内容更直接地指向一些影响人一生的、稳定的人格特质（Bouchard，1994，Loehlin，1992）。这些研究和其他一些研究的结果表明，人们在外倾—内倾（开朗大方或腼腆内向）、神经质（承受高度焦虑或具有偏激的感情反应的倾向）和自觉性（个人的干练、负责任和考虑问题的全面程度）等特性上的变异可以更多地（65%）以遗传差异而非环境因素来解释。

当然在科学界，并非所有人都愿意从表面意义上接受这些研究结果。对鲍查德和莱肯研究的批评意见体现在多个方面（见Billings et al.，1992）。有些评论家声称，这些研究者并没有尽可能完整地公布他们的研究数据，因此，不能孤立地对他们的研究结果进行评价。他们还认为，有很多研究报告表明鲍查德和莱肯没能考虑到的一些环境因素对双胞胎确实有重大影响。

此外，一些研究者对双生子研究的“相同环境假设”（equal environment assumption）方面提出了批评（例如Joseph，2002）。这些争论认为鲍查德和莱肯有关遗传影响的许多结论都源于这样一个假

设：共同抚养的同卵双生子和异卵双生子的生长环境是相同的。这些批评家认为这个假设是无效的，与同卵双生子相比，异卵双生子受到的对待是非常不同的。他们对使用双生子这种研究方法去研究基因影响因素提出了质疑。然而，也有一些文章反对这种批评，并且支持“相同环境假设”（例如，Kendler et al.，1993）。

近期应用

在1999年，鲍查德评价了明尼苏达双胞胎研究档案上所有有关“先天—后天”的研究证据（Bouchard，1999）。他总结道，从整体上看，在人格中40%的变异和智力中，50%的变异都以遗传为基础。他又反复强调了前面已经讨论过的观点，即基因驱使你对环境作出选择，或驱使你回避某种特定的环境和行为。

在明尼苏达双胞胎中心进行的研究至今仍很活跃。其中有些饶有趣味的新研究已将人的复杂特性和行为，比如爱情、离婚和死亡等作为研究对象，而以前很少有人会认为这些特性受到遗传的影响（见明尼苏达双胞胎家庭研究，2007）。他们还将“坠入爱河”的人们作为研究对象，研究人们对梦中情人或白马王子的选择是否也同样受遗传的预先安排。结果发现，事实并非如此！然而，研究者们确实发现了一种与离婚有关的基因，以及一种与死亡年龄有关的基因。

最后，鲍查德和莱肯的研究被应用到人类克隆问题的哲学讨论上（见Agar，2003）。如果一个人被成功地克隆了，问题是，正如你可能会想到的，他或她的基本特质，比如独特的人格，在多大程度上会被传递到克隆体上？那些反对克隆的普遍论据是，担心人类的身份可能会因此被改变、贬低或者丢失。另一方面，双生子研究的结果，诸如鲍查德和莱肯的观点，表明“在某种情况下，克隆人从某种程度来说是生命的延续……然而……通过克隆延续生命必须有助于防止对克隆技术的滥用，而不仅仅是许不许可的问题（Agar，2003，P.9）。另一项独立研究探讨了同卵双生子和克隆问题（Prainsack & Spector，2006），研究者发现，同卵双生子很少考虑自己作为同卵双生子在实际生活经历中的遗传因素。此外，从个人视角来看，虽然他们认为克隆人是一种不自然的或不道德的观点，但是更需要关注有关人类克隆的伦理上的根本原因。当然，这更像是哲学上的讨论，而不是基因的讨论，但是它产生的有趣结论，丰富了我们的思想。

Agar, N.(2003).Cloning and identity.Journal of Medicine and Philosophy, 28, 9-26.

Arvey, R., Bouchard, T., Segal, N., & Abraham, L.(1989).Job satisfaction: Environmental and genetic components.Journal of Applied Psychology, 74(2)，187-195.

Arvey, R., McCall, B., Bouchard, T., & Taubman, P.(1994).Genetic influences on job satisfaction and work value.Personality and Individual Differences，17(1), 21-33.

Billings, P., Beckwith, J., & Alper, J.(1992).The genetic analysis of human behavior: A new era? Social Science and Medicine, 35(3), 227-238.

Bouchard, T.(1994).Genes, environment and personality.Science, 264(5166)，1700-1702.

Bouchard, T.(1999).Genes, environment, and personality.In S.Ceci et al.(Eds.), The nature-nurture debate: The essential readings, pp.97-103.Malden, MA: Blackwell.

Joseph, J.(2002).Twin studies in psychiatry and psychology: Science or pseudoscience? Psychiatric Quarterly, 73, 71-82.

Kendler K., Neale M., Kessler R., Heath A., Eaves L.(1993).A test of the equal environment assumption in twin studies of psychiatric illness.Behavioral Genetics, 23, 21-27.

Loehlin, J.(1992).Genes and environment in personality development.Newbury Park, CA: Sage Publications.

Minnesota Twin Family Study (2007).What's special about twins to science? Retrieved March 10, 2007 from http://www.psych.umn.edu/psylabs/mtfs/special.htm.

Prainsack, B., & Spector, T.D.(2006).Twins: A cloning experience.Social Science & Medicine, 63(10), 2739-2752.

研究4 小心视崖

Gibson, E.J., & Walk, R.D.(1960).The “visual cliff.” Scientific American, 202(4), 67-71.

在心理学界，人们常常提起一则有关S.B.先生（为了保护他的隐私，用字母代替真名）的趣事。S.B.天生看不见东西，直到他52岁时，一种新出现的先进手术（即现在已很普遍的角膜移植术）才使他恢复了视力。虽然S.B.获得了视力，但并不意味着他自动地获得了和我们正常人一样的视知觉能力。这一点在手术后不久，他的视力还没完全恢复的时候，就变得十分明显。当S.B.从医院的窗户向外望时，他十分好奇地发现楼下的地面上一些小东西在移动着，于是他爬到窗边，想探下身子看个究竟。幸亏医院的工作人员及时阻止了他的行动。他当时住在四楼，下面移动着的小东西是汽车！尽管S.B.现在能够看见东西，但他不能知觉深度。

实验心理学家感兴趣的领域之一是我们用视觉感知和解释周围世界的能力。由此产生了一个重要的问题——这种能力是天生的还是后天习得的——先天—后天的问题再一次被讨论。许多心理学家认为，我们最重要的视觉能力就是深度知觉。你可以想象一下，一个人如果没有深度知觉能力，其生存将变得多么困难和痛苦。你会撞上东西、不能判断猛兽离你有多远，或者会从悬崖上掉下来。因此，从逻辑上可以假设人对深度的知觉能力是一种天生存在的机制，是不需要经验来发展的。然而，正像埃里阿诺·吉布森（Eleanor Gibson）和理查德·沃克（Richard Walk）在他们文章中写到的那样：“婴儿在爬行和蹒跚学步阶段，很容易从或高或低的地方跌下来。大人们必须很警惕，以防止婴儿爬到儿童床的床栏边或楼梯口。当婴儿的肌肉协调能力成熟后，他们开始能自己避免跌落。通常人们认为，孩子是从经验中学会认识容易摔跤的地方的，即从有过自己跌落和受伤的经验中学习”（P.64）。

在实验室里，研究者欲科学地研究深度视知觉能力。为此，他们构思并设计出了一种叫作“视崖”的实验装置。

理论假设

如果你想找到动物或人在发展过程中获得深度知觉的关键点，一种方法是把他们放在悬崖边上，看他们能否让自己不掉下去。这个建议十分可笑，因为从伦理上说，这可能伤害到无法知觉深度（特别是高度）的被试。“视崖”则解决了这个问题，因为它能造成视觉上的视崖，而实

际上并没有真正的悬崖存在。他们是如何做的，我们将在后面详细介绍。这种装置的重要意义在于，可以把婴儿或小动物放在视崖上，观察他们是否能知觉到这种悬崖并进行躲避。如果他们不能这样做，而从“悬崖”上跌下来，他们也并没有真正落下的危险。

吉布森和沃克在这篇文章中持有“先天论”的观点，他们相信深度知觉和避免从高处跌落的能力是自动生成的，是我们生理机制的一部分，因此它们不是经验的产物。经验主义者则持相反的观点，认为这种能力是习得的。吉布森和沃克的视崖允许他们提出这样的问题：人或动物在发展的哪个阶段才能对深度和高度刺激作出有效的反应？对不同种类和生存环境不同的动物，这种反应出现的时间是否相同？

方法

视崖装置的组成：一张4英尺[1]高的桌子，顶部是一块透明的厚玻璃（见图1—6和图1—7）。

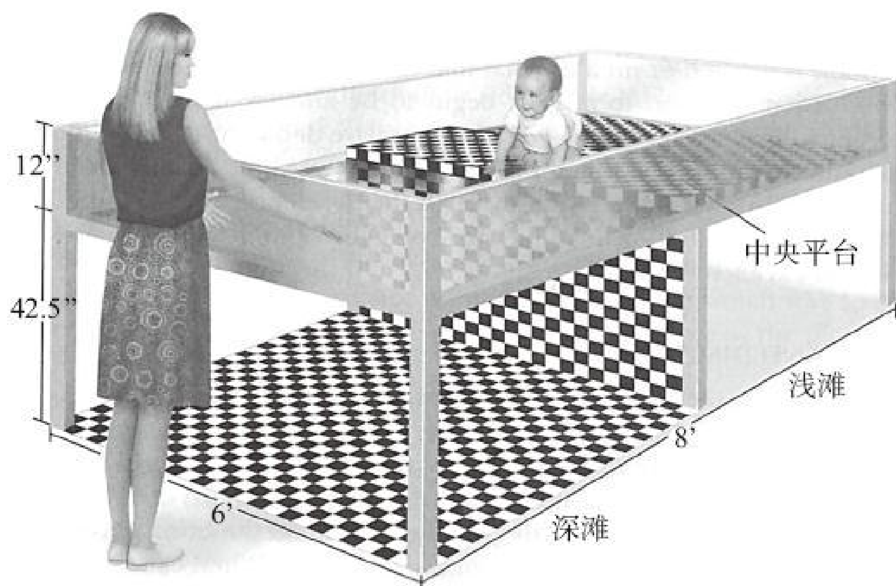


图1—6 吉布森和沃克的视崖



图1—7 视崖测试情景

资料来源：引自Mark Richards/Pohto Edit Inc.。

桌子的一半（浅滩）是用红白格图案组成的结实桌面。另一半是同样的图案，但它在桌子下面的地板上（深渊）。因此，从浅滩边上看，图案好像直接落到了地面，但实际上一块玻璃覆盖着整个桌面。在浅滩和深渊的中间是一块1英尺宽的中间板。使用这种装置对婴儿施测的过程十分简单。

这项研究的被试是36名年龄在6到14个月之间的婴儿。这些婴儿的母亲也参加了实验。每个婴儿都被放在视崖的中间板上，先让母亲在深的一侧呼唤自己的孩子，然后再在浅的一侧呼唤自己的孩子。

为了比较人类与小动物的深度知觉能力，对其他种类的动物也进行了视崖实验（当然没有母亲的召唤）。这些动物被放在视崖的中间“地带”，观察它们是否能区别浅滩和深渊，以避免摔下“悬崖”。你可以想象一下，将许多种小动物汇集到康奈尔大学的心理实验室做实验，是一个多么独特而有趣的情景。这些动物包括小鸡、小海龟、小老鼠、小绵羊、小山羊、小猪、小猫和小狗。有人可能会问：“这些动物是否是同一天接受测试的？”

请注意，这项研究的目的是检测深度知觉是后天习得的还是天生的。这个实验方法之所以巧妙，是因为可以回答或至少开始回答这个问题。毕竟，我们无法向婴儿或小动物询问他们是否知觉到深度，而且就

像上面提到过的那样，他们也不能在真正的悬崖上进行试验。心理学上很多问题是随着新的实验方法的进步而得到答案的。吉布森和沃克早期研究的结果在这方面给我们提供了一个很好的案例。

结果和讨论

在研究中，9名婴儿拒绝离开中间板。虽然研究者没有解释这个问题，但这可能是因为婴儿太过固执。当另外27位母亲在浅的一侧呼唤她们的孩子时，所有的孩子都爬下中央板并穿过玻璃。然而只有3名婴儿，当母亲在深的一侧呼唤他们时，极为犹豫地爬过视崖的边缘。当母亲从视崖的深渊一侧呼唤孩子时，大部分婴儿拒绝穿过视崖，他们远离母亲爬向浅的一侧；或因不能够到母亲那里而大哭起来。婴儿已经意识到视崖深度的存在，这一点几乎是毫无疑问的。“通常他们能透过深的一侧玻璃注视下面的深渊，然后再爬向浅滩。一些婴儿用手拍打玻璃，虽然这种触觉使他们确信玻璃的坚固，但还是拒绝爬过去”（P.64）。

这些研究结果能证明人类知觉深度的能力是天生的而不是后天习得的吗？明显不能！因为在这项研究中，所有婴儿至少已经有了六个月的生活经历，在这段时间内，他们可能通过尝试错误而学会了知觉深度。然而，六个月以下的婴儿由于不具备自主运动的能力，所以不能接受实验。这也是吉布森和沃克用各种动物作为实验参照的原因。众所周知，大部分非人类的动物获得自主运动的能力比人类婴儿要早得多。动物测试的结果很有趣——不同种类动物知觉深度能力的发展与它们的生存需要有关。

例如，小鸡出壳后，就必须马上开始自己觅食。当研究者把出生不足24小时的小鸡放在视崖上接受测试时，它们从不会犯跌下深渊的错误。

小山羊和小绵羊在出生后很快就可以站立、行走。从能站立的那一刻起，它们对视崖的反应和小鸡一样准确而可预测，一次错误也没有。当研究者把一只出生仅一天的小山羊放在深渊的一侧玻璃板上时，它变得惊恐呆滞，表现出防御性姿态。随后，如果把它推向浅滩的一侧，它则变得轻松自在，并跳向看似坚实的表面。这说明视觉起着完全的控制作用，动物虽然能感觉到在深渊一侧的上面有坚实的玻璃，但这种感觉没能影响它们的反应。

而对于小老鼠则是另一回事了。它们对浅滩没有表现出明显的偏好。你认为这种差异为什么会存在呢？在你得出“小老鼠比较愚蠢”这样的结论以前，先来看看吉布森和沃克的解释：小老鼠对视觉依赖性不大。实际上，它们的视觉系统不够发达。因为它们属于夜间活动的动物，在黑暗中靠鼻子上坚硬的触须帮助其四处活动，靠嗅觉感知食物的方位。所以，当一只小老鼠被放在中间地带时，它不会被视崖所欺骗，

因为它不是用视觉来决定走哪条路的。对小老鼠的触须而言，深侧和浅侧的玻璃在感觉上没有区别，所以小老鼠离开中间地带走向深侧的概率与浅侧是相同的。

你可能会猜测，小猫也有同样的结果。它们主要在夜间活动，而且也有触须。然而，猫是食肉动物，不像老鼠一样是食腐动物。因而它们更加依赖视觉。相应地，自从它们能自主运动起（约在出生后4周），小猫便具有极好的深度知觉能力。

尽管有时这篇研究文章（特别是讨论部分）听起来好像是给孩子讲的动物故事，但还要说明的是，各种动物中在视崖上成绩最差的是海龟。研究者选择的海龟是水栖类动物，因为研究者猜测由于海龟自然生活的环境是水，它们可能更喜欢深的一侧。然而，事实证明，聪明的海龟知道它们并不是真的在水里，它们中的76%都爬到浅的一侧。但是也有24%的小海龟“越过边界”。“数量相对较多的小海龟选择了深侧，这一事实可能表明这种海龟的深度知觉能力比其他动物要差，也可能是它的自然生活环境使它较少‘害怕’跌落的状态”（P.67）。很明显，如果你生活在水中，在防止跌落方面，深度知觉能力对于个体生存的价值将会减小。

吉布森和沃克指出，他们所有的观察结果和进化论完全一致。也就是说，所有种类的动物，如果它们要生存，就必须在能够独立行动时发展感知深度的能力。对人类来说，这种能力到6个月左右才会出现，但是对于鸡、羊来说，这种能力几乎是一出生（一天之内）就出现了；而对于老鼠、猫和狗来说，大约在4周时出现这种能力。因此，作者得出结论：这种能力是天生的。因为如果通过尝试错误而获得这种能力可能会带来过多潜在的、致命的危险。

那么，如果说我们在生理上已经具备了足够的能力，为什么仍有那么多儿童会摔跤呢？吉布森和沃克解释说，这是因为婴儿的深度知觉能力比他们的行动技能成熟得早。实验中，许多婴儿在中间带转身时，会借助深侧玻璃来转换支撑点，当他们开始转向浅侧并爬向母亲时，有的甚至倒在深的一侧。如果那儿没有玻璃，一些婴儿会真的摔下悬崖。

批评与后续研究

对于研究者所得结论，批评得最多的一点是它们是否能证明人类天生具备深度知觉。前面已经提到，当婴儿在视崖上接受测试时，他们已经学会了避免这种情况。后来，有研究把年龄在2到5个月之间更小的婴儿放在视崖深侧的玻璃上，这时所有的婴儿都表现出心率变慢。这种心率变慢是感兴趣的信号，而不是恐惧的信号，恐惧应伴随着心率加快（Campos et al.，1978）。这表明这些婴儿还没有习得对落差的恐惧，稍后他们将学会这种规避行为。这些发现与吉布森和沃克的结论恰恰相

反。

然而，应该注意到，对于我们什么时候开始能感知深度这一问题，学术界仍在继续争论（先天论者对经验论者），许多为寻找这一问题答案的研究都使用了吉布森和沃克的视崖装置。此外，另一使用视崖的相关研究得到了一些有趣的结果。

由索兹等人（Sorce et al.，1985）进行的研究是个很好的例子。他们把1岁的婴儿放在视崖上，落差不深也不浅，大约30英寸[2]。当婴儿爬向视崖时，他会停下来往下看。在另一边，与吉布森和沃克的研究一样，母亲在那里等待。有时母亲根据指令在脸上做出害怕的表情，有时则看起来兴高采烈和兴趣盎然。当婴儿看到母亲害怕的表情时，他们会拒绝再向前爬。然而，当看到母亲高兴的表情时，大部分婴儿会再次检查悬崖并爬过去。当落差被变得很浅时，婴儿不再观察母亲的表情而径直向前爬。婴儿这种通过非语言交流以改变行为的方式叫作“社会参照”。

近期应用

吉布森和沃克发明的视崖装置，在如今研究人类发展、认知、情绪甚至心理健康等方面都发挥着重要的影响。

伯杰和爱波勒（Berger & Adolph，2003）最近的一项研究引用了吉布森和沃克的早期研究，他们关注初学走路的儿童是如何分析深度特征，特别是在过桥任务中。研究者哄着年龄较小的初学走路的儿童（16个月），走过不同宽度的桥，一些桥有扶手，而另一些没有。研究者发现，婴儿显然更有可能从较宽的桥穿过，而不是较窄的（16个月的婴儿非常聪明）。然而，更有趣的是，研究者发现，如果在有扶手的情况下，婴儿更有可能试着从窄的桥穿过。“在迈上桥之前，婴儿会试探桥和扶手，并且设计出另外的过桥策略的婴儿，更有可能成功穿过桥。”这些对传统的工具概念化提出了挑战：婴儿使用扶手作为保持平衡的手段，并执行在没有扶手的情况下不可能完成的目标导向任务”（P.594）。

如何使用技术性视崖的思想来帮助恐高症（一种非理性的恐惧高度）的人？在人群中，恐高症的流行率是1/20。假设你是一名治疗师，如果你想让恐高症患者“直面他们对于高度的恐惧”的话，你也许不会有太多客户了。为什么？因为他们实在太恐惧了！当一个人确实患有某种恐惧症，他会不惜一切地去避免接触害怕对象或情境。结合吉布森和沃克的一项研究发现，能引起恐高症者害怕的虚拟跌落，其治疗效果对患者的作用很小。研究提示我们考虑更有效、更便捷的治疗手段，来减少非理性的恐惧（Coelho et al.，2009）。

结论

由于吉布森和沃克的创造发明，行为科学家能够用一种明确、系统的方式去研究深度知觉能力。行为科学家对这种能力和其他认知能力到底是天生的还是后天习得的这一问题一直处于争论中，真相可能是一种折中的情况，即认为是先天与后天的交互作用所致。正如很多研究者指出的，深度知觉能力可能在婴儿一出生就已经具备了，但害怕跌落和避免危险是婴儿到了能够爬行的年龄并遇到危险后，通过经验习得的。

但是不管怎么说，像视崖这种优秀的方法，是一种研究方法学上的进步，它使我们得以继续去探寻答案。

Berger, S., & Adolph, K.(2003).Infants use handrails as tools in a locomotor task.Developmental Psychology, 39, 594-605.

Campos, J., Hiatt, S., Ramsay, D., Henderson, C., & Svejda, M.(1978).The emergence of fear on the visual cliff.In M.Lewis & L.A.Rosenblum (Eds.), The development of affect.New York: Plenum Press.

Coelho, C., Waters, A., Hine, T., & Wallis, G.92009).The use of virtual reality in acrophobia research and treatment.Journal of Anxiety Disorders, 23, 563-574.

Sorce, J., Emde, R., Campos, J., & Klinnert, M.(1985).Maternal emotion signaling: Its effect on the visual cliff behavior of 1-year-olds.Developmental Psychology, 21 ，195-200.

注释

[1]1英尺≈30.48厘米。——译者注

[2]1英寸≈2.54厘米。——译者注

第二章 知觉与意识

心理学家之所以对知觉与意识的研究非常感兴趣，是因为它们阐述并展现了人的心理与环境的相互影响。思考一下，每时每刻由周围的各种刺激组合而成的数以百万计的信息，对你的感官进行着怎样的狂轰滥炸？你的大脑不可能对所有的信息进行加工。所以，你的大脑就要把这些大量来自感官的信息资料组织成具有一定形式和意义的单元。这就是心理学家所说的“知觉”。

显然，你的意识水平，通常也称为意识状态，在很大程度上控制你对什么进行知觉以及大脑怎样组织这些信息。当你度过白天、黑夜、一周、一年甚至一生时，你会体验到各种意识状态的变化：集中注意力（或不集中）、做白日梦、胡思乱想、进入睡眠或开始做梦，你可能在某一刻被催眠，也有可能服用某种精神类药物（甚至是咖啡因和尼古丁等）。这些都是你的意识状态的变化，它们改变着你对周围世界的感知，进而影响你的行为。

在知觉和意识的研究领域内，一些极具影响力又饶有趣味的研究都集中在儿童早期的知觉能力、睡眠、梦和催眠上。本章将以一项影响深远的著名研究作为开始，该研究为研究者提供了一种精妙而卓越的研究方法，可用于研究语前儿童甚至是出生几天的儿童的思维过程和知觉过程。这种方法称之为偏好注视，有助于人们了解婴儿的大脑功能以及他们是如何对世界进行概念化的；本章的第二篇研究报告包含了两项改变心理学的研究，因为它们（1）发现了REM睡眠（快速眼动睡眠）并（2）揭示了快速眼动睡眠和梦之间的关系；第三篇研究报告论述了一项非常有影响也极受争议的研究。它提出，梦并不像弗洛伊德和其他研究者所说的那样，是来自于人潜意识中的神秘信息，而是纯粹由睡眠时大脑中随机的生理电刺激所致；第四项研究是影响传统心理学思想的众多研究之一。该研究反对一种为人普遍接受的观点，即催眠是一种强有力的独特意识状态。该研究表明，被催眠的人与清醒的人没有什么不同，只不过前者更易被驱动罢了。

研究5 注视偏好

Fantz, R.L.(1961).The origin of form perception.Scientific American, 204(May), 61-72

如果你想知道别人对周围世界的知觉，一个很简单的办法就是去问他们。通常情况下，他们都会回答你。但是，你试过问婴儿这些问题吗？由于不会说话，几乎所有的婴儿都没有办法告诉你他们的想法和认识。即使能告诉你，也不会有很多，并且最有可能的是，他们完全不理解你的问题。

如果你有机会与婴儿接触（每个人都有不同程度的机会），你可能常常想：“我很好奇这个婴儿在想什么！”或者“如果这个婴儿能说话该多好……”但不幸的是，这不可能发生，约翰·特拉沃尔塔（John Travolta）的《飞越童真》（Look Who’s Talking）[1]系列电影除外。但是，纵观心理学的历史，对于婴儿的研究始终处于一个非常重要的位置（仅本书中就包含了七个专注于婴儿的研究）。

在本章中，我们将会讨论罗伯特·范茨（Robert L.Fantz）关于下面问题的一些发现：“我们如何研究婴儿的认知过程？”“我们如何真正地窥视婴儿的大脑正在干什么？知觉了什么？以及理解了多少？”

20世纪50年代，来自克里夫兰西储大学（现为凯斯西储大学）的心理学家罗伯特·L·范茨，注意到了关于婴儿的一些很有趣的事情，但他观察的对象不是人类婴儿，而是刚孵出的小鸡。范茨报告说，小鸡几乎是从破壳而出时就能够很好地知觉周围环境，并马上开始搜索和啄食（见“小心视崖”一节，以了解更多关于小鸡的认知能力）。在某种程度上，这就意味着小鸡确实有着比人类婴儿更优秀的知觉能力，这使小鸡成为该问题的理想研究对象。但需要强调的是，当心理学家研究动物时，他们的终极目标是将研究结果应用于理解人类行为，我们将在后面进一步讨论这个问题。

理论观点

范茨之前的研究者已经证明，人类婴儿有一些初步知觉周围世界的方法，如趋向光亮、区分基本颜色、察觉运动等能力。但是，正如范茨所指出的：“婴儿不能对形状、图案、大小或体积等刺激作出反应已经引起了众多争论；总之，婴儿不能知觉形态。”（P.66）范茨对该观点表示怀疑。因此，在20世纪50年代晚期到20世纪60年代早期，他着手开发

了一种新技术，以便研究者能获得婴儿知觉过程的更多细节，探查知觉能力发展的时间进程以及判定婴儿知觉技能的复杂性。他提出人类婴儿在刚出生时，并非与初生的小鸡完全不同，他们能够知觉不同形态；我们可以通过观察婴儿如何“分析”他们的世界（即他们“看什么”以及“看多久”）来证明这件事。这种被称为“偏好注视”的研究婴儿心理的方法，在心理学界得到了迅速的推广，并延续至今，带来了一场理解婴儿心理研究的革命。

方法

范茨通过一个简单的实验，证明了新生小鸡能知觉什么、不能知觉什么。范茨在小鸡没有任何啄取食物的经验之前，向其呈现不同形状、大小的物体，并记录小鸡啄每个物体的次数。与锥形物体相比，小鸡更多地啄圆形物体；圆形物体比三角形物体啄的多；球形物体比扁平物体啄的多，并且面对大小不同的圆形物体，小鸡更喜欢直径比1/8英寸大一点或小一点的物体。先前没有任何学习经验，小鸡就能知觉形态——很明显，它们喜欢那些更像食物（种子或谷物）的形状。

范茨在其文章中写到，你现在可能想到的是：“当然，适用于鸟类的规律，未必适用于人类。”（P.67）他考虑到了这种可能性：鸟类知觉形态（多种鸟类都有）的这种先天能力可能在灵长类动物（包括人类）进化过程中没有得到发展，或者，也许灵长类动物在出生后，需要经过一段时间的发展或学习，才能习得这些能力。所以，当范茨将注意力转回到灵长类婴儿时，需要一个新的研究方法，很显然这是因为灵长类婴儿不会去“啄取”任何食物，并且即便有这种倾向，其所需的动作也未充分发育（当然，婴儿不会这样做的原因是，他们并不特别喜欢谷物和种子）。

然而，婴儿有一种与小鸡啄食相似的行为可以作为观察的指标：他们会凝视物体。如果范茨能够用一种方法发现婴儿对一些形态物体的注视比对另一些物体的注视多或时间长，其唯一的解释就是他们能区分出差异，即能知觉形态。研究从与人类基因最相似的黑猩猩婴儿开始，范茨和他的同事们发明了所谓的“可见密室”（looking chamber），大致是一个放置于大盒子里的舒适摇篮。在盒子顶端有两个开口，用于给婴儿呈现物体，还有一个窥视孔，用于研究者观察婴儿的观看行为。当研究者确定与其他物体相比，黑猩猩婴儿对特定的物体表现出系统性偏好（由凝视时间决定）时，他们开始运用相似的程序来研究人类婴儿。

研究者没有干涉婴儿平常的安排或活动，只是将婴儿放入舒适的观察箱，给他们呈现各种各样成对的物体。婴儿的年龄范围在1周~15周。呈现给婴儿的刺激包括：立体的、有纹理的圆盘；球体；椭圆形人脸；包含杂乱的人类面部特征的椭圆形以及多种复杂性的形状和图案（见图2—1）。研究者将这些物体进行多次两两配对，观察并记录每个

1分钟的试次内，婴儿注视每个物体的总时间以及在每组物体中，他们“喜欢”（注视时间长）哪个。他们的实验结果强有力地证明了：所有年龄段的婴儿都拥有知觉和辨别不同形态的能力。

结果

在实验的第一个阶段，研究者给婴儿看多组不同的黑白测试图对，包括：一个有靶状图的正方形和一个有横纹的正方形；一个棋盘和一个扁平的没有图案的正方形；一个加粗的十字和一个圆；还有一对完全相同的三角形作为对照刺激。结果如图2—1所示，婴儿明显“喜欢”更复杂的形态（靶状图、条纹和棋盘），且这种偏好具有跨年龄的一致性。这表明，区别这些形态的能力是先天的，出生时即可表现出来。大约8周开始，相比于条纹，婴儿更喜欢靶心；相比于扁平的正方形，婴儿更喜欢棋盘。这种时间上的延迟意味着婴儿出生后的2个月可能发生了学习，或者说脑和/或视觉系统的成熟导致了这些变化。

与这些发现一样有兴趣的是，婴儿的能力和刚孵化出的小鸡研究之间的一个重要的联系仍不明确。如果人类婴儿出生时就具备无需学习的、天生的区分形态的能力，我们必须要问这是为什么。对于小鸡而言，答案是相当直接明了的：它们辨别形态，以便寻找食物以求得生存。对于人类婴儿而言，知觉特殊形态的先天能力如何具有生存价值呢？也许原因是相似的。范茨写道：

> 人对婴儿的重要性就像谷物对小鸡的重要性。对于从其他物体中区分并识别一个人而言，面孔形状是人类最具特点的方面……因此，婴儿对面孔图片所产生的选择性认知也是毫无意外的（P.70）。

换句话说，人类婴儿不用依赖形态知觉来寻找食物和生存，而是依赖他人的照料。就像小鸡可以很好地识别特定形态那样，婴儿在知觉过程中偏好人类面孔也就不难理解了。事实也的确如此。

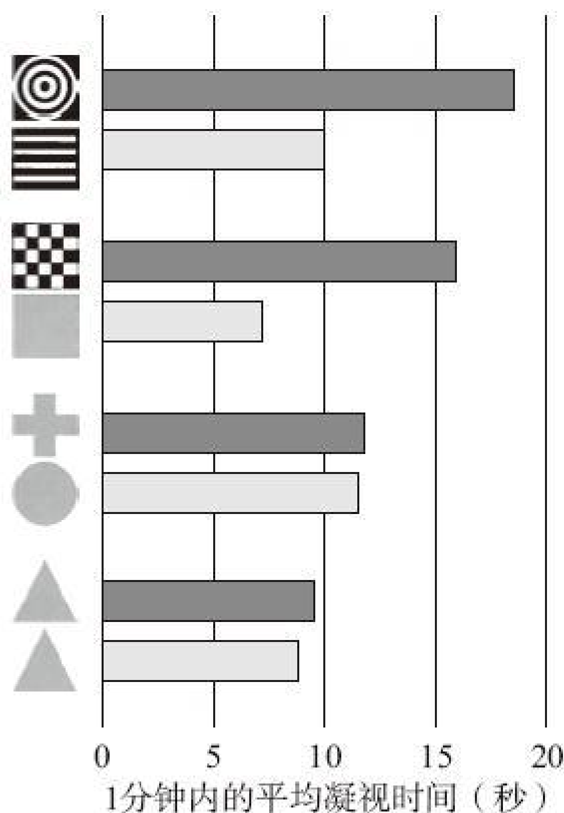


图2—1 220次测验中，婴儿对形态对的兴趣作为平均注视时间的函数

资料来源：Fantz，1961，P.70。

范茨的团队给49个婴儿（4天~6个月）呈现3个大小一样的椭圆盘。一个画有面孔特征；一个有面孔特征，但被打乱；第三个为控制盘，是一个在一段有黑色填充的椭圆盘，该椭圆盘的黑色面积与其他的两个椭圆盘黑色面积相同（如图2—2）。

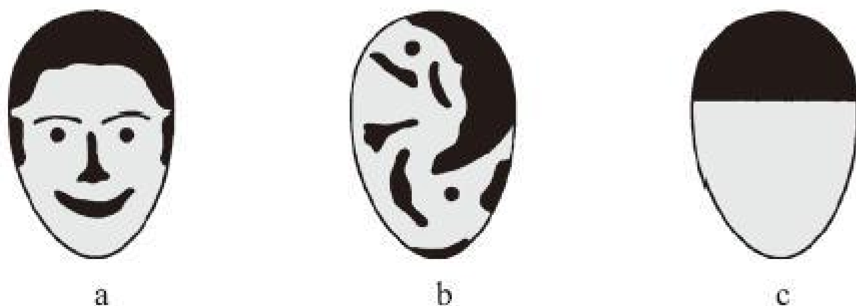


图2—2 范茨的面孔图测验

资料来源：Fantz，1961 P.72。

婴儿对面孔特征的椭圆盘表现出有较高的兴趣，与图2—2中的b相比，婴儿喜欢a；与c相比，更喜欢a和b，并且专心地凝视它们，几乎忽视了控制盘。所有年龄段的婴儿的这种偏好强度都相同，再次证明了基本形态知觉在出生时即有，而排除学习或发展因素。

在这篇文章最后的研究报告中，研究者再次验证了人类婴儿知觉面孔图片的能力。呈现给婴儿6个直径为6英寸的圆盘，设计如下：（1）面孔；（2）靶状图；（3）打印页面（如报纸或者课本）的随机片段；（4）全红色的；（5）全荧光黄色的；（6）乳白色的。记录婴儿第一次看每个圆盘的时间。你认为他们看哪个盘子看的时间最长呢？如果你说“面孔”，那你就猜对了。与任何其他形态或颜色相比，他们更多地凝视面孔盘（见图2—3）。

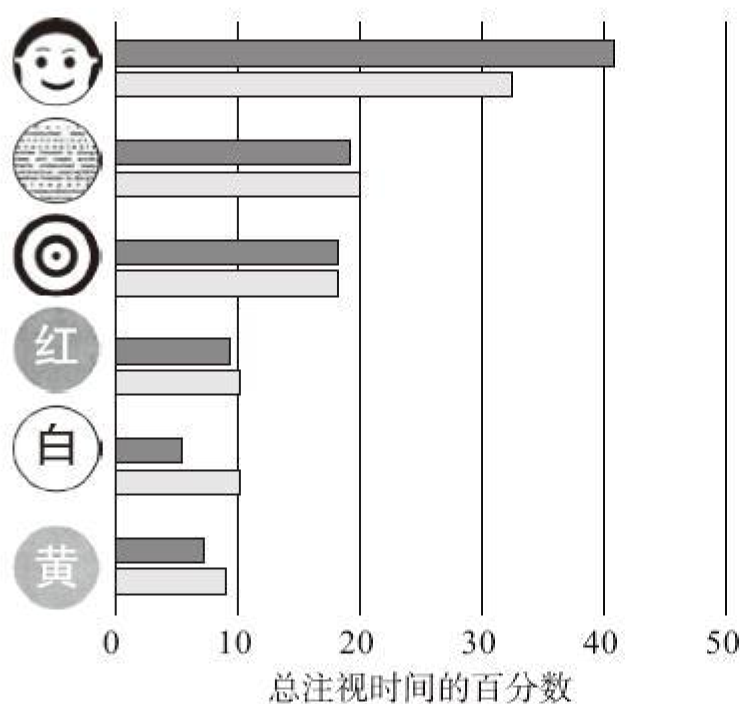


图2—3 婴儿对图形和颜色的注视时间

注：黑条：8~12个月；灰条：年龄大于12个月。

后续研究和最新应用

和这本书里的很多研究一样，该研究明显地改变了心理学。这里有两个原因：开创性的发现以及研究者为使这些发现成为可能而发明的研

究方法。直到20世纪中期前，许多行为学和生物医学研究者都假定，婴儿出生时几乎没有任何知觉和感觉能力，只有随着个体与周围环境的互动，才能发展或习得到全部或大部分的能力。心理学上的“空白”新生儿的观点，接受起来相对简单，因为那时我们并没有可以揭示初生婴儿真实能力的研究方法。范茨提供了偏好注视（preferential-looking）的方法，打开了一扇相当准确的通向婴儿心理的门。这种方法如今被普遍地使用，它对心理学的意义就像显微镜之于生物学：当想要研究婴儿如何思考时，该方法成为研究者首选的工具之一。当然，婴儿带着大量的知觉技能来到这个世界的发现并不会降低学习和发展的重要性。但是，使用范茨的方法发现的先天技能似乎为婴儿未来的生存和成长提供了一个平台。正如范茨指出的：

> 环境的固有知识由新生小鸡对似乎可吃的形态的偏好决定，且被婴儿对形态种类的兴趣决定，这种形态将在以后帮助物体辨别、社会回应和空间定向。这种原始的知识为大量的通过经验形成的知识的累计奠定了基础（P.72）。

范茨的发现激起了婴儿知觉能力的研究革命。你可以看到范茨的方法的独创性对发展和认知心理学领域的影响。例如，在婴儿认知领域，一些世界领导级研究者，如伊利诺伊州大学的婴儿认知实验室的芮妮·贝拉吉恩（Renee Baillargeon）和哈佛大学发展研究实验室的伊丽莎白·斯佩克（Elizabeth Spelke），在许多研究（见Talbot，2006，作为这项研究的回顾）中大量运用了范茨的偏好注视的研究方法。此外，范茨的研究也有利于说明在吉布森和沃克的经典视崖实验（见第1章）已经详细探讨的儿童深度知觉的出现时间和发展进程。

也许对于范茨的研究最重要的延伸，要归于堪萨斯大学的弗朗西斯·霍洛维茨（Frances Horowitz）。他发现，除了偏好注视，婴儿会对一遍一遍看相同的刺激感到厌烦（Horowitz，& Paden et al.，1972）。当你给婴儿呈现一个新的视觉图片（例如在范茨的研究中使用的图片）时，他们会在一段时间内注视它，但如果相同的刺激重复呈现时，婴儿的注视时间会减少，这就叫作习惯化（habituation）。如果变换图片，他们似乎又有了兴趣，且注视时间变长，这种反应叫作去习惯化（dishabituation）。通过结合优先观察、习惯化和去习惯化的方法论，研究者能够了解婴儿，甚至是新生儿所“知道”的周围世界的更多细节。

例如，在最近的一项研究中，研究者想要看何时人们能够习得区分“可能”物体和“不可能”物体的能力（Shuwarai，Albert，& Johnson，2007）。你无疑已经理解了所谓的不可能物体，就是我们经常称为错觉的东西。图2—4列举了可能和不可能物体间的区别。你会对不可能物体注视时间长，对吗？婴儿也是如此。运用偏好注视和注视时间的方法，研究者发现4个月大的婴儿注视不可能物体的时间较长，表明他们能意识到这种差异，就好像是说：“我能够发现这个物体有错误，且需要尝试着指出它来！”

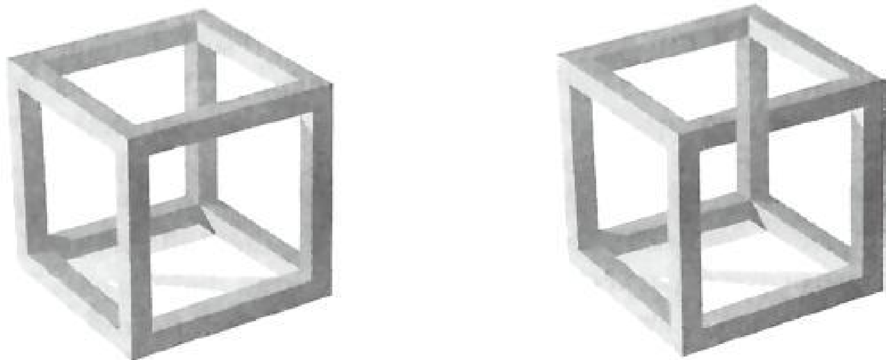


图2—4 4个月的婴儿可以区分可能的和不可能的物体

这只是每年由发展心理学家和其他行为科学家所做的成百上千个研究中的一个例子，他们最基本的方法依赖于罗伯特·范茨的发现。这些方法允许我们对婴儿的大脑活动一窥究竟，因为以前从来没有了解过婴儿的知觉和想法。事实上，每当我们再看一次，就会发现他们是“精明者”，因为他们能知觉到比我们预期的要多的东西。

Horowitz, F.D., Paden , L., Bhana, K., & Self, P.(1972).An infant-controlled procedure for studying infant visual fixations.Developmental Psychology, 7, 90.

Shuwairi, s., Albert, M., & Johnson, S.(2007).Discrimination of possible and impossible objects infancy.Psychological Science ，18(4), 303-307.

Talbot, M.(2006, September 4).The baby lab.The New Yorker, 82(27), 91-101.

注释

[1]这是一部美国家庭喜剧系列电影，影片中的婴儿被戏剧化地赋予了成人的思维和语言能力。——译者注

研究6 睡眠，毫无疑问就会做梦

Aserinsky, E., & Kleitman, N.(1953).Regularly occurring periods of eye mobility and concomitant phenomena during sleep.Science，118, 273-274.

Dement, W.(1960).The effect of dream deprivation.Science，131，1705-1707.

正如你所看到的，这一部分与其他部分相比有些不同，因为在这里我们探讨了两项研究。第一项研究发现了与睡眠和做梦有关的一些基本现象，为第二项研究提供了可能性。本章的主要内容是维廉·迪蒙特（William Dement）关于做梦剥夺的研究，但为了让读者对此有所准备，我们首先介绍尤金·阿瑟瑞斯基（Eugene Aserinsky）的发现。

1952年，当阿瑟瑞斯基还是个研究生时，就开始了有关睡眠的研究。他研究的一部分是观察睡眠中的婴儿。他发现，当婴儿睡眠时，总会周期性地出现主动性眼动，而在其他时间里，仅偶尔出现慢速眼球运动。他认为可能这些主动性的眼动会与做梦有关。然而，婴儿不可能告诉他自己是否在做梦。所以，为了验证他的想法，他将研究范围扩展到成人。

阿瑟瑞斯基和他的同事纳撒尼尔·克特曼（Nathaniel Kleitman）征召了20名正常成年人作为被试。电极装置被连接在被试眼部周围的肌肉上，电极的另一端用导线与灵敏的电子测量仪器相连接，导线一直延伸到隔壁房间，研究者在那儿监控被试的睡眠。然后被试就可以正常地入睡了（每个被试参加观测不止一夜）。整个夜晚，无论在主动性的眼动期，还是很少或根本没有眼动的时期，主试很可能会把被试叫醒，并对其进行“询问”。叫醒被试的目的是询问他们是否在做梦以及是否还记得梦的内容。研究结果相当具有启发性。

综合所有被试的观察结果，在快速眼动睡眠期间，被试总共被唤醒27人次。其中，20人详细报告了非常形象的梦境，其余的7人报告“感觉到在做梦”，但无法回忆详细内容。在无眼动期间，共唤醒被试23人次，其中，有19人次没有报告任何梦境；其余4人次模糊地感到好像是在做梦，但无法描述其梦境的内容。在有些情况下，允许被试整夜睡眠而不被打扰。结果发现，在平均7个小时的睡眠过程中，被试出现3到4次主动眼动期。

今天我们已经非常熟悉阿瑟瑞斯基对REM睡眠（快速眼动睡眠）或者有梦睡眠的发现，而在当时却并不是很引人注目。他的发现引出了大量关于睡眠和做梦的研究，而且其数量与日俱增。多年来，随着生理记录仪器和研究方法的日趋成熟，我们可以进一步深化阿瑟瑞斯基的发现，并揭开睡眠的神秘面纱。

例如，我们现在已经知道，当你入睡后，睡眠共经历四个阶段：开始是浅睡眠阶段（阶段1），接着进入越来越深的睡眠阶段。然后，当你达到最深的深度睡眠阶段（阶段4）后，便开始依次返回到最初的阶段，你的睡眠会越来越浅。当你即将再次进入睡眠的第一阶段时，就会出现一个叫作REM睡眠的特殊阶段。做梦大多出现在REM睡眠阶段。然而，与普遍想法相反的是，科学研究发现，在REM睡眠期间，人们并不经常挪动身体。来自大脑的电化学信息能够麻痹你的肌肉，使你的身体不能动弹。有了这样一种生存机制，人们才不会用肢体去演绎梦境，从而避免了伤害自己或者出现更糟的后果。

短暂的REM睡眠期后，你又重新开始睡眠的四个阶段，称为“非快速眼动睡眠”（NON-REM或NREM）。整个夜晚，REM睡眠和NREM睡眠交替出现约5到6次（大约入睡90分钟以后第一次出现REM睡眠），随后NREM睡眠越来越短，而REM睡眠越来越长（因此，这就是人在清晨时做梦较多的原因）。顺便提一下，每个人都会做梦。虽然有少部分人声称他们从不记得做过梦，但研究证明所有人都会做梦。

所有的这些认识都产生于阿瑟瑞斯基在20世纪50年代早期对REM睡眠的研究发现。继阿瑟瑞斯基之后，斯坦福大学的威廉·迪蒙特（William Dement）为我们提供了很多关于睡眠和梦的非常有价值的信息。大约从阿瑟瑞斯基的发现开始，迪蒙特就开始了对睡眠与做梦数十年极具开创性的研究。

理论假设

在前人的研究中，有一项发现对迪蒙特影响最大，那就是每个人每天晚上都会做梦。正如迪蒙特在他的文章中所说的：“由于睡觉的人每晚都会出现大量的梦境，你可能会问，从某种意义上说，大量的梦是否成为人的生命所必不可少且至关重要的一部分。”（P.1705）进而，他提出，如果这一前提成立，那么“梦被完全地或部分地剥夺了，人是否还能继续正常地活动？梦在心理学意义上是必需的，还是在生理学意义上是必需的，或是二者兼而有之”（P.1705）。

迪蒙特决定通过剥夺被试做梦的机会来尝试回答这几个问题。起初，他尝试使用具有镇静作用的药物来阻止被试做梦，但药物本身对被试睡眠模式的影响太大，以至于无法得到有效的结果。所以，他决定使用“有些过激的方法”，就是每当被试在夜里进入REM睡眠中，就把他们叫醒。

方法

这篇论文报告了一项研究计划中前8名被试的情况，该项研究与睡眠和做梦有关，且现在仍在进行。被试全部为男性，年龄从23岁到32岁。被试在他们通常的入睡时间来到睡眠实验室。为记录脑电波图形和眼球运动情况，主试把小电极连接到被试的头皮和眼部周围。与阿瑟瑞斯基的研究一样，这些电极的导线穿墙而过连到隔壁房间里，这样被试可以在安静而黑暗的房间中入睡。

研究程序是这样的：开始的几个晚上，允许被试整夜正常地睡眠。这样做是为了确定每个被试在正常情况下做梦数量的基线和整个的睡眠模式。获得这些资料之后，下一步就是要剥夺被试REM睡眠或有梦的睡眠。在以后的几个夜晚（不同被试被连续剥夺REM睡眠的数量从3夜到7夜不等），每当电极传导的信息表明被试开始做梦时，实验者就唤醒被试，要求他在床上坐起来直至有迹象表明他已经完全清醒，几分钟后才允许他们再次入睡。

迪蒙特提到一个要点，即在此研究中，主试要求被试在除研究以外的其他任何时间都不能睡觉。这是因为如果被试睡着了或打了个盹儿，他们就可能会做梦，这将会污染实验结果。

在完成做梦剥夺的研究阶段后，被试就进入了实验的“恢复阶段”。在恢复阶段（1至6夜不等），允许被试整夜安然入睡而不被打扰。继续像以前那样，对他们这一阶段的梦进行电子监控并记录做梦的数量。

之后，给所有被试放几天假（毫无疑问，他们肯定会很开心）。接下来他们中有6人回到实验室，继续进行另一系列干扰睡眠的研究。这次实验中唤醒的夜晚数及每晚唤醒的次数“完全重复了做梦剥夺实验阶段的情况。唯一不同的是唤醒被试的时间，即改为在两次眼动（做梦）之间的一段时间内唤醒被试。在做梦开始时，允许被试一直睡下去而不被打扰，当梦自然地结束以后再来唤醒他们”（P.1706）。最后，就像在做梦剥夺阶段随后所做的那样，给被试相同数量的“恢复夜晚”，这被称为“控制恢复”阶段。这一阶段的目的是为了消除这样一种可能，即无论被试做梦与否，剥夺做梦的效果不能简单归因于每晚被叫醒的次数。

结果

表2—1总结了研究的主要发现。在建立基线的夜晚，当被试不被打扰地睡觉时，他们每晚的平均睡眠时间为6小时50分钟。被试用于做梦的平均时间为80分钟，或占整个睡眠时间的19.5%（见表2—1第1列）。迪蒙特从开头几夜的结果中发现，不同被试之间用于做梦的时间惊人地相似。事实上，其做梦时间的变化仅在正负7分钟之内！

表2—1做梦剥夺实验的研究结果

被试	1 基线夜做梦时间百分比	2 做梦剥夺夜晚的数量	3a 唤醒次数 第一夜	3b 唤醒次数 最后一夜	4 恢复期做梦时间百分比	5 控制阶段做梦时间百分比
1	19.5	5	8	14	34.0	15.6
2	18.8	7	7	24	34.2	22.7
3	19.5	5	11	30	17.8	20.2
4	18.6	5	7	23	26.3	18.8
5	19.3	5	10	20	29.5	26.3
6	20.8	4	13	20	29.0	—
7	17.9	4	22	30	19.8(28.1)*	16.8
8	20.8	3	9	13	—**	—
平均数	19.5	4.38	11	22	26.6	20.1

*恢复阶段第二夜；

**被试在恢复阶段之前退出研究。

该研究的主要目的是检验做梦剥夺或REM睡眠剥夺所产生的影响。与此有关的第一个发现是：在做梦剥夺阶段，阻止被试进入REM睡眠所需要的唤醒次数。正如你在表2—1中（3a列）所看到的，第一夜，实验者为了阻止被试进入REM睡眠，唤醒被试的次数在7至22次之间。然而，随着实验的进程，为了阻止被试做梦而唤醒被试的次数越来越多。在最后一个夜晚，唤醒被试的次数达到13至30次（3b列）。平均来看，在做梦剥夺结束的那一夜，被试试图做梦的次数增加到最初的2倍。

第二个也可能是最能说明问题的发现，被试在经历了几个被阻止做梦的夜晚以后，其做梦的时间明显增加。表2—1（第4列）中的数字反映了恢复阶段第一夜的情况。这一夜所有人做梦时间平均为112分钟，占整个睡眠时间的26.6%（而第1列基线夜的做梦时间平均为80分钟，占整个睡眠的19.5%）。迪蒙特指出，有两名被试的REM睡眠时间没有表现出显著的增加（被试3和被试7）。如果将他们排除后，再来计算全部做梦时间，其平均值为127分钟，占整个睡眠的29%。这比基线夜的平均值增加了50%。

虽然在表2—1中，只报告了第一个恢复夜晚的数据，需要注意的是，大多数被试在接连5个夜晚连续表现出做梦时间的增加（与基线相比）。

“等一下！”你会想。也许这里做梦时间的增加与剥夺REM睡眠根本无关。也许这只是由于这些被试被唤醒得过于频繁。那么，迪蒙特会让你记住，他已为你敏捷的观察做了充分准备。被试中，有6个人在休息了几天以后又回到实验室，并严格地重复了实验过程，只不过是在两次REM睡眠期间被唤醒（次数相同）。在这种条件下，没有出现做梦时间的显著增加。在控制唤醒阶段以后，用于做梦的平均时间为88分钟，占整个睡眠的20.1%（第5列）。与第1列的80分钟和19.5%相比，未发现

显著差异。

讨论

迪蒙特从以上研究结果中试探性地得出这样的结论：人需要梦。倘若不允许人们做梦，经过连续几夜剥夺做梦的睡眠，似乎就会增加人们做梦的压力。在经历做梦剥夺的阶段后，被试做梦的数量有所增加（比较3a列和3b列），且做梦的时间也显著延长（比较第4列和第1列），这两点在他的发现中显而易见。他还指出，这种增长会持续几个晚上，以便在数量上尽量补偿被剥夺的梦。尽管迪蒙特那时没有使用这一专业术语，如今这个重要的发现被称为“REM反弹”（REM-rebound）效应。

在这篇简明而不同凡响的论文中还有一些有趣的额外发现。如果重新查看一下表2—1，就会发现，如我们先前所说，有2名被试未表现出显著的REM反弹（被试3和被试7）。在研究中被试相对较少时，对这些特例的解释就显得十分重要了。迪蒙特发现，不难解释被试7做梦时间的增加较少这一现象：“他在恢复期的第一天晚上没能表现出做梦时间的增加，是由于来实验室之前，他在一个聚会上喝了几杯鸡尾酒，因此预期会出现的做梦时间的增加被酒精的抑制作用抵销了。”（P.1706）

然而，对被试3解释起来就比较困难。尽管在做梦剥夺阶段，他的唤醒次数增幅最大（从7变为30），但在恢复期的5天晚上，他没有表现出任何REM反弹。迪蒙特承认这名被试是其研究的特例。他从理论上推测认为，这名被试可能具有异常稳定的睡眠模式，以至于可以抵御任何变化。

最后，对8名被试进行监控，以观察是否会因为REM睡眠的剥夺而使其行为受到影响。所有被试在REM睡眠干扰阶段都出现了轻微的焦虑、烦躁和注意力不易集中的症状。有5名被试报告说在做梦剥夺阶段食欲明显增加，其中的3名被试体重增加了3至5磅[1]。在控制唤醒阶段，这些行为症状并未有所表现。

研究发现的意义及后续研究

在迪蒙特开创性研究以后的40多年，我们掌握了大量有关睡眠和做梦的知识。其中一些已在本章前几部分进行了简要的讨论。我们知道，迪蒙特1960年的论文中绝大部分内容经受住了时间的考验。所有人都做梦，而且如果我们某个晚上因为某种原因而无法做梦的话，我们会在下一晚做更多的梦。的确，在我们做梦的需要中有某些本能的东西。事实上，在很多动物身上都可以观察到REM反弹效应。

有一个仅仅被迪蒙特作为轶事来报告的偶然发现，现在看来却意义重大。剥夺人REM睡眠的方式之一是通过使用酒精制品或其他药物，如安非他明、巴比妥酸盐等。这些药物在增加沉睡趋势的同时，它们会抑制REM睡眠并使人在夜晚的大部分时间里都保持在深度的NREM睡眠阶段。正是由于这个原因，很多人无法戒除为了睡眠而服用安眠药或酒精制品的习惯。一旦他们停止使用这些东西，过于强烈的REM反弹效应会妨碍他们的睡眠，导致他们害怕睡觉并重新开始服用这些药物来避免做梦。这种问题的一个极端事例是一些嗜酒成瘾的人们，很多年来，他们可能一直在剥夺自己的REM睡眠。一旦他们停止饮酒，REM反弹作用极其强烈，以至于当他们清醒时也会出现做梦的现象！这也许可以解释为因酒精中毒而引起的震颤性谵妄现象（delirium tremens，DTs），这种病常常会出现可怕的幻觉（Greenberg & Perlman，1967）。

迪蒙特用几十年的时间继续他早期的开创性研究，关注着剥夺做梦对人行为的影响。在后来的研究中，他在更长的一段时间内剥夺被试的REM睡眠，但未发现有任何证据证明这种剥夺引起了有害的变化。他总结道：“十年的研究未能证明大量的不良影响是由剥夺REM睡眠引起的，即使这种剥夺是长期而有选择的。”（Dement，1974）

源于迪蒙特这项早期研究的其他研究报告认为，在REM睡眠期间，大脑中蛋白质的合成比NREM睡眠期间更甚。也有人认为，这些化学变化可能意味着人们将新信息整合入大脑记忆结构的过程，甚至可能成为人格变化的生物基础（Rossi，1973）。

近期应用

如今，研究睡眠和梦这一领域的专家们已经普遍接受了“阿瑟瑞斯基发现REM睡眠”这一事实。大多数与睡眠、做梦或睡眠紊乱有关的研究都是以他的实验研究为基础的。因此，他与克特曼的早期研究经常被新近的科学论文频繁引用。

迪蒙特对阿瑟瑞斯基研究的扩展研究，更为研究睡眠模式的论文所广泛引用。一项最近的研究发现，人们在NREM睡眠中做梦的次数可能比我们预想的要多（Suzuki，et al.，2004）。研究人员让被试白天打盹，这样通常能比晚上睡眠更快进入NREM睡眠。研究发现，当询问被试打盹时做梦的情况时，只有那些真正进入到NREM睡眠状态的被试才回答经常做梦。然而，研究人员还发现，“在NREM阶段打盹的被试的做梦情况，在数量、生动性和情感等方面都不如REM阶段打盹的做梦情况”（P.1486）。

另一项研究与迪蒙特和阿瑟瑞斯基的基本研究主张相关，认为在REM睡眠阶段，人们会发展出一种前意识，即意识发展所需要的一种大脑的基本组织形式（Hobson，2009）。这种大脑基本形式的发展从出

生前一直持续到整个儿童时期。霍布森在研究中提出，早期的REM睡眠为我们提供了一个现实世界的虚拟模型，这可以帮助我们完成日常生活中的各种任务。这个理论可以解释两个现象：为什么婴儿比成人会有更多的REM睡眠？为何人类的大脑每晚总是维持最少数量的REM睡眠？

结论

迪蒙特这位一直致力于睡眠医疗项目的研究专家，于2000年在斯坦福大学出版了他的学术著作《睡眠的价值：睡眠医学领域中对健康、幸福和高质量睡眠间重要关系的开创性探索》（The Promise of Sleep: A Pioneer in Sleep Medicine Explores the Vital Connection Between Health, Happiness and a Good Night’s Sleep）。这本书是为非科学家的普通人所著。迪蒙特花费四十年时间研究睡眠，积累了丰富的知识，并利用这些知识帮助我们了解高质量睡眠的重要性，以及如何去改善睡眠质量。在这本书中，迪蒙特（2004）将我们描述成一个“不健康睡眠的群体”，并以睡眠研究者的身份提出了自己的目标：

> 在我职业生涯的大部分时间里……我之所以在不停地改变社会对睡眠的应对方式，是因为当前的方式，或者根本不成为一种方式，是如此地糟糕……对于数以百万甚至数十亿的人来说，如果他们能够理解一些简单的原则，其生活质量就能得以提高，然而他们却不这样做。一想到这些，我感到非常伤心。改变社会及其组织机构应对睡眠的方式，是我所能够想到的最有意义的一件事，当然，也是我一直以来努力去实现的目标（P.4-5）。

要想更多地了解迪蒙特在斯坦福大学的人类睡眠研究中心所做的出色工作，访问http://med.stanford.edu/school/psychiatry/humansleep。

Dement, W.C.(1974).Some must watch while some must sleep.San Francisco, CA: Freeman.

Dement, W.C.(2000).The promise of sleep: A pioneer in sleep medicine explores the vital connection between health, happiness and a good night’s sleep.New York: Dell.

Greenberg, R., & Perlman, C.(1967).Delirium tremens and dreaming.American Journal of Psychiatry，124，133-142.

Hobson, J.(2009).REM sleep and dreaming: Towards a theory of protoconsciousness.Neuroscience，10, 803-813.

Rossi, E.I.(1973).The dream protein hypothesis.American Journal of Psychiatry，130，1094-1097.

Suzuki, H., Uchiyama, M., & Tagaya, H., at al.(2004).Dreaming

during non-rapid eye movement sleep in the absence of prior rapid eye movement sleep.Sleep, 27(8)，1486-90.

注释

[1]1磅≈0.454千克。——译者注

研究7 类别天成

Rosch, Eleanor.H.(1973).Natural categories.Cognitive psychology, 4, 328-350.

在1934年拍摄的电影《起立欢呼》（Stand Up and Cheer）中，伟大的演员和舞者秀兰·邓波儿所扮演的丝蒂宾·芬奇（Stepin Fetchit）坐在一幢老建筑的走廊台阶上，检视着自己的一双老旧的鞋子，并充满哲学意味地思考着："鞋子为什么叫作鞋子呢？"她扮演的剧中人物总想知道某事物为什么叫现在的名字。许多心理学家也在通过各种方式来寻求这个问题的答案。行为科学家主要通过关注认知（思维）和知觉（人类对周围世界的解释）来对这个问题进行研究。

这些研究领域的一个基础性问题就是概念（concepts）。概念是对所经历事物的一种心理表征，根据事物的一些共同特征对其进行分类（例如家具、蔬菜、动物、专业、鞋子等）。由于可以将一组物体进行分类，因而概念对于提高信息加工的效率来说是非常有用的。例如，你知道某件特定家具是椅子，因为它符合你对于椅子的"概念"。因此，每当你见到式样不熟悉的椅子时，你不需要重新进行学习，就能知道这个特定物件也被称为椅子，因为它处于你对椅子的这一分类中。

因为我们在之前的讨论中提到了椅子，你现在可能正在思考椅子，对么？什么特征符合你的"椅子概念"呢？你可能认为椅子有腿、有椅面、有椅背可以靠。即使有些椅子不完全符合这些特征（比如躺椅和摇椅没有椅腿），它们仍然符合你对椅子的分类。然而，如果你见到豆袋椅（bean bag，俗称懒人沙发），而且不知道它是什么，你可能不会把它称为椅子。实际上，你可能根本不知道它叫什么。

认知心理学家最感兴趣的问题是：关于物体分类的知识是从哪儿获得的？在1970年之前，被普遍接受的传统或"经典"观点认为，分类是我们所使用语言的功能。换句话说，类别的存在是由于我们有相对应的词汇。例如，我们关于一种动物的分类是可以下蛋、能飞、有羽毛、能鸣叫，这种类别的动物我们称之为"鸟"。这种传统观点认为，如果我们没有关于鸟的词汇，这种类别或概念就不会存在。

因此，概念和分类由于语言的变化在不同的文化间也应该发生变化，而且有相关的例子。一个经常被引用的例子是生活在北极附近的因纽特人的语言中有12个关于"雪"的词汇，而在英语当中只有1到2个。很明显，由于所生活的环境的气候，因纽特人在关于雪的交流有更大的灵活性，这反映在他们的语言当中。在南太平洋岛上生活的人，在他们的语言当中根本就没有关于雪的词汇，因此科学家认为对他们来说没有关于雪的概念。

许多年来，关于概念起源的这种观点被一些社会科学领域的科学家，例如心理学家、人类学家、语言学家和社会学家认为是理所当然的。在20世纪70年代早期，加利福尼亚大学伯克利分校的埃莉诺·罗施（Eleanor Rosch）发表了一些研究结果，对这个观点提出了挑战，并且彻底地颠覆了分类领域认知心理学的研究。她提出，分类并不是来自于语言，而是天然存在的，与人类知觉的生物能力有关。

在这里，我们将要呈现的具有里程碑式意义的研究包括两个独立的实验以及一些技术规程。为了更清楚地说明，以及受篇幅的限制，我们只会对研究中的第一个实验进行详细的说明。

理论观点

罗施认为，如果之前被普遍认可的观点是正确的，那么属于同一类别的所有物体在该类别中应该具有大致相同的状态——也就是它们同等程度地符合这一类别。然而，她观察到，事实并非如此。相反，人们往往认为一个类别当中的部分“样例”，相比其他某些样例会更符合这一类别（她将这些样例称之为原型）。举个例子来说，考虑一下“鸟”这个类别，然后在你脑中迅速形成鸟的图像。你首先可能想到的是知更鸟、冠蓝鸟、鹪鹩或者雀类（可能是一只乌鸦或者一只鹰）。你可能不会首先想到鹅、鸡、鸵鸟或者企鹅。根据罗施的观点，这是因为知更鸟比鸡更符合你关于鸟的原型（或者说理想的例子）。换句话说，知更鸟符合你全部或者大部分关于鸟这一类别的描述，因此你认为知更鸟更加符合鸟的分类。相反，鸵鸟符合鸟的特征比较少——不能飞，也不能鸣叫，而且体型很大。因此，对于大多数人来说，鸵鸟并不是关于鸟的一个原型。

罗施认为，对于一个物体是否符合某一类别，大部分情况下并不存在明显的界限，我们的心理类别的边界是很模糊的。我们通过将物体与原型进行比较来确定其是否属于某一个类别。同时，她认为就算某种语言没有关于某一类别的词汇，那么这一类别仍然是存在的，并且具有心理现实性。

为了检验这一理论，在20世纪60年代后期，罗施去到了位于新几内亚的丹尼人（Dani）部落（参见本书第22项研究中艾克曼对这一国家文化的讨论）。丹尼人迄今为止仍处于石器时代，而且他们的语言中不包含一些现代文化已有的概念。我们将要讨论的一项罗施早期研究，关注了颜色的分类。在英语当中，颜色主要有11个类别：红色、黄色、绿色、蓝色、黑色、灰色、白色、紫色、橙色、粉色和棕色。研究已经确定了说英语者对这些类别的“焦点颜色”：对每一种颜色类别最好的例子（或者说颜色原型）。例如，说英语者认为消防车的红色是红色这一颜色类别的焦点色（也可以说是最像红色）。这种红色相比于其他的“淡红色”或者红色，在被判断是否属于红色时更快，也更简单。

然而，丹尼人关于颜色只有两个分类：深色冷色调的“mili”和明亮暖色调的“mola”。罗施认为，如果分类是由语言决定的，这一“经典”观点是正确的，那么丹尼人应该只能够加工两种颜色概念。然而，罗施认为，所有人都有多种关于颜色的分类，而不仅仅是两种。而且，这种分类是与生俱来的，在进化的过程中已经植入到人的脑中。

为了检验这一观点，她决定教给丹尼人八种或为焦点颜色或为非焦点色颜色的新词。她假设存在多种颜色类别而不仅仅是两种颜色类别，对于丹尼人来说具有心理现实性，即使他们语言当中并没有关于这些类别的命名词汇。如果这一观点是正确的，那么对于丹尼人来说，焦点颜色（或者说原型色）要比非焦点颜色学起来的更快，也更加简单。

方法

被试

罗施研究中的被试是年轻的男性丹尼人，确定这些被试中没有色盲。对他们的颜色知识进行测试，确定他们关于颜色的词汇知识仅知道“mili”（深色）和“mola”（明亮色）。有趣的是，丹尼人并不记录他们的年龄，因此根据他们的身材和身体成熟度，研究者判断这些被试的年龄大致在12到15岁之间。这些被试自愿参加这项研究，当然会给予这些被试报酬。

在完成颜色学习阶段的程序后，这些被试被分为12人的小组。这里我们将要讨论两个最重要的实验组。

颜色刺激

在学习阶段使用的刺激材料是具有光泽的颜色盘（color chips），类似于当你需要粉刷房屋时在油漆店所见到的。

然而，这些颜色盘是根据科学方法制成的，代表了准确波长下的颜色。实验中使用的颜色类别包括粉色、红色、黄色、橙色、棕色、绿色、蓝色和紫色。对于其中一组被试呈现的颜色是焦点色，特定颜色的原型（消防车的红色）。罗施认为这些颜色被认为是普遍代表自然形成的类别，因此，无论在何种语言和文化背景下，这些颜色都是容易识别的。

对于另一组被试，八种颜色盘是处于两种焦点色之间的颜色，对这些颜色说英语的人可能会称它们为“红–棕色”、“黄–绿色”。这些颜色被命名为模糊或者“非焦点色”。

程序

罗施及其同事所面临的第一个挑战是如何使用丹尼人的语言给这些颜色命名。这个问题并不像听起来那么简单，因为为了避免使用词汇本身所造成的偏差，所使用的词汇对于这些被试来说应该有相同的使用频率，同等的熟悉性，以及相同的意义。罗施发现丹尼人对于他们的近亲有多种称谓。罗施所说的近亲是类似于宗族的家庭群体。这些称谓符合她所要寻找的颜色替代词的要求，因而被用来表征不同的颜色类别，让被试学习。为了避免偏好或偏见，一个被试如果恰好有相应的亲人，那么对于这个被试则不使用这个词汇来命名颜色。

每一个被试被告知他们将要学习一种新的语言，并由实验者教给他们。在第一天开始的时候，给被试呈现每种颜色及其相应的名称，实验者读，被试重复。之后，将这些颜色顺序打乱，重新呈现给被试。如果被试能正确地命名颜色，实验者会给予表扬，如果命名错误，会被告知正确的名称。在被试学习5到12天后，会对被试颜色命名的学习结果进行测试，并对被试学习进展进行记录，直到所有的被试都能对颜色进行正确的命名。

在完成学习阶段的任务之后，所有被试会进行下一个任务，检验被试这种新的一般类别概念的能力是否真正的形成，能否向新情境迁移，还是仅仅局限于所学习过的特定颜色。

为了检验这个问题，实验者会给被试呈现多种颜色，其中包括八种没有学习过的颜色，询问他们这些颜色是否属于某种颜色类别。计算两组被试的迁移任务成功率（用来说明之前的学习能否迁移到不同的情境）

结果

如果人能够天生地知觉颜色的类别（就像罗施所假设的那样），那么在颜色学习任务中，那些学习焦点颜色的丹尼人被试要比那些学习非焦点颜色的被试学得更快，因为焦点色更符合颜色概念的原型。图2—5概括说明了两组被试在学习阶段的学习进展。

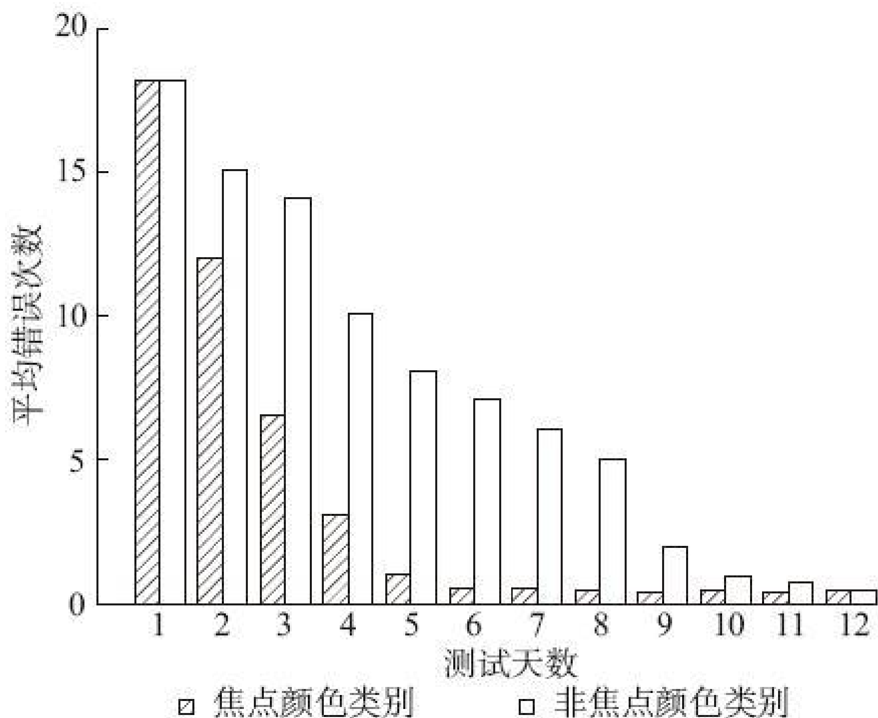


图2—5 学习焦点颜色和非焦点颜色类别学习的平均错误率

学习颜色原型组被试整体的错误率为8.54，而学习模糊颜色组被试整体的错误率为18.96。两组之间的差异是非常显著的。如果仔细地观察图2—5，能够发现学习原型颜色的被试只需要5天就能很好学习颜色类别，而学习非原型颜色的被试则需要11天。

罗施认为，说明通过这种学习方式获得的颜色类别的能力迁移到的新情境是非常重要的。这不仅仅是对特定颜色的学习，而是形成了一种“有用”的概念。在对没有学习过的颜色被询问是否属于某个特定类别的任务中，如果被试不能将学习过的颜色进行迁移，那么被试的正确率应该在猜测水平的12%。罗施的研究结果表明被试回答的正确率大约为90%。

罗施还报告了一个额外的信息，在学习阶段完成之前，在学习非焦点色的被试中，有4名被试感觉非常沮丧，要求退出实验。实验者做了很大的努力才说服被试继续完成实验任务。在学习焦点颜色的被试组中并没有遇到这个问题，他们都非常乐意完成实验任务。

为了获得一个相对简单的结果，看起来仍然有很长的路要走。但是，正如在本节开头所提到的，罗施的这项研究结果和其他研究，以及其他人的一些研究对于理解我们的大脑是如何工作的产生了深远影响。接下来，我们会首先概括性地介绍罗施对本研究结果的讨论，然后对该研究之后出现的大量相关研究进行简单的介绍。

讨论

罗施发现了一种测量方法，可以用来检验看起来几乎不能检验的理论。你能想象一种物体的概念或者类别在英语（或者其他语言）当中并不存在而恰好在丹尼人的语言中正好没有关于颜色的分类？可能存在，但是很难发现，而测量起来会更加困难。研究不存在颜色类别知识的文化本身就是一项非常精妙的研究。但是她的贡献更多地源自她的发现。

这项研究的主要结果是，即使在没有颜色知识文化中的人也能够学习颜色类别知识，并且通过原型学习颜色要比通过非原型学习颜色更快。这些发现说明，特定的概念存在于我们的脑中，不管使用的语言是什么，也不管是否使用过这些概念，这就是主要的发现。因为这些概念看起来是人类生物组织的一部分，罗施称之为“自然类别”（她的文章题目）。这一项研究之所以对心理学有如此大的影响，是因为它彻底地颠覆了原有观念。之前普遍认为语言产生了概念，而这种观点不仅仅受到了挑战，而且与之前的观点相反，语言中的概念是围绕着自然存在的类别而形成和发展的。

罗施进行了总结，并指出这项研究更深层的有意义：

> 简而言之，本研究所呈现的关于颜色类别的学习及其结构的意义可能远远超出颜色领域本身：（a）在其他领域中可能也存在类似的自然类别；（b）甚至在非知觉领域，人工原型（非知觉类别的最佳样例）一旦获得，可能会以类似自然类别原型的方式影响该领域的学习与加工（P.349）。

这就意味着，物体（自然的或人工的）是通过将其与相应类别的原型相比较，是否符合原型来分析和分类的，而不是通过审视其是否符合一般的语言定义来判断是否属于某个类别。

后续研究

在罗施的早期研究之后，大量的研究都支持自然类别的存在以及类别原型在概念的形成过程中的作用。罗施及其同事，还有其他的一些研究者证实了她的这项早期研究的发现，并展示了这项研究的深远意义。

例如，正如我们之前提及的，根据严格的语言学定义，物体的类别之间应该是有明确的界限的，然而罗施认为概念之间并不存在清晰、明确的界限。相反，概念之间的界限是模糊的，甚至会存在一定程度的重叠（Rosch，1975）。如果我们回过头来再看一下关于“鸟”这个概念的例子，你认为鹬鸵（kiwi）是鸟类么？那么蝙蝠是鸟类么？你可能有关于鹬鸵的知识，知道它属于鸟类（即使它不能飞，也不鸣叫，蹲在树上），但是当你想到鸟的时候，鹬鸵通常不会出现在你的脑海（不过现在可能会了）。

另一方面，你知道蝙蝠不是鸟，即使它能飞，发出特定的鸣叫声，甚至有的蝙蝠生活在树上。正是由于蝙蝠的这些特征，一些人在一定意义上持有蝙蝠是鸟的观点。再举一个例子，考虑一下“水果”这个类别，你首先想到的是哪些水果？人们首先想到的往往是苹果、桔子或者香蕉。那么，番茄是不是水果呢？番茄可能是水果，但它是一个非典型的样例（与你的原型相比），因为番茄与你心目中的水果原型相距较远。请记住，心理学家所关心的是，人是如何思考的，而不是实际上是否正确（顺便说一下，kiwi[1]对于鸟这个类别来说，不是一个典型的样例，对于水果来说也不是一个典型的样例）。

近期应用

罗施对新几内亚丹尼人的研究表明自然类别的存在之后，多种研究方法用来揭示人类是如何对周围的世界概念化的（参见罗施1978年的完整讨论）。其中一种方法是，让被试对列出的物体是否属于某个类别进行等级评定（例如1到10的等级）。对于“狗”这一类别，德国的牧羊犬可能被评定为10，而法国斗牛犬可能被评定为3（这与狗的品种没有关系，而是人们认为在多大程度上评定它是狗），如图2—6所示。其他的一些研究使用反应时间作为指标测量一些事物是否属于某个心理类别（Dovidio，1986；Rosch & Mervis，1975；Unyk，1990）。



图2—6 法国斗牛犬

注：众所周知，法国斗牛犬属于狗，但当你被问到狗的类别时，你很少会想到它。这是因为斗牛犬并不符合你对狗的类别判定（除非你自己养了一只）。

还有一种方法，让被试听一句描述，例如“火鸡是鸟”，然后尽快地按键，作出是或否的反应。使用这种研究方法得出的结果表明，样例与

类别的原型越近，那么判断得越快。对“火鸡是鸟”的判断反应要明显地慢于“知更鸟是鸟”的判断。

第三种方法是，让被试列出或画出某个类别的样例。在给定的时间内，被试通常将那些对于某个类别更具代表性的样例列出。例如，让你画出家具这个类别，在你画出碗架、书柜之前，你通常可能先会画出椅子、沙发或桌子；或者让你列出人类的情绪，你可能会首先列出高兴、生气，之后才会想到困惑、愤怒（Fehr & Russell，1984）。

和本书当中其他的研究类似，罗施关于自然类别和原型的发现改变了心理学对于概念的认识。在过去的40年中，一些研究扩展了她的研究结果或者对此提出了质疑。例如，一些研究发现，虽然罗施的原型理论看起来是有效的，但是并不意味着彻底放弃了概念严格的语言学定义。在2010年的一项研究中，虽然承认我们关于罗施的讨论，但也报告了另外一些研究，包括对巴布亚岛另外一个亚文化群体被试的研究，发现对于概念的学习看起来并不是基于概念的原型，而依赖的是语言的线索（Tylen et al.，2010）。

事实看起来可能是当我们需要概念的准确描述时，我们会使用概念的语言学定义。之前提到的水果类别可以作为一个很好的例子。如果有人问你：“你想要一个水果吗？”你不会想到你脑中关于水果的原型，例如苹果、桔子或香蕉。如果当你回答了“是的”而给了你一个松果的时候（虽然松果是松树的果实），你可能会表现得很惊讶。然而，在特定的情境下，水果准确的语言定义也是有用的。例如，当你在大自然中漫步的时候，遇到了不常见的植物，并且结了奇怪的果实。即使这些果实与你脑海中水果原型一点都不相似，而你关于水果正式的语言学定义可能会让你说出：“这株植物结了水果。”

Dovidio, J.(1984).Concept of emotion viewed from a prototype perspective.Journal of Experimental Psychology: General，113, 464 - 486.dio:10.1037/0096-3445.113.3.464.

Fehr, B., & Russell, J.(1986).Racial stereotypes: The contents of their cognitive representations.Journal of Experimental Social Psychology, 22, 22-37.

Rosch, E., & Mervis, C.B.(1975).Family resemblances: Studies in the internal structure of categories.Cognitive Psychology, 7, 573-605.

Rosch, E.H (1975).Cognitive representations of semantic categories.Journal of Experimental Psychology: General，104，192-233.

Rosch, E.H (1978).Principles of categorization.In E.Rosch & B.Lloyd (Eds.), Cognition and Categorization.NJ: Lawrence Erlbaum.

Tylen, E., Weed, E., Wallentin, M., Roepstorff, A., & Frith, C. (2010).Language as a tool for interacting minds.Mind & Language, 25, 3-29.

Unyk, A.M.(1990).An information-processing analysis of expectancy in music cognition.Psychomusicology: A Journal of Research in Music Cognition, 9, 229-240.

注释

[1]kiwi在英语中既可以指鹬鸵，也可以指猕猴桃。——译者注

研究8 行动，如同被催眠了一样

Spanos, N.P.(1982).Hypnotic behavior: A cognitive, social, psychological perspective.Research Communications in Psychology, Psychiatry, and Behavior, 7，199-213.

我们都非常清楚意识状态的改变与睡眠和梦有关。另一种与意识状态改变有关的现象就是催眠。催眠常常被人们认为是对人的心理有控制作用的神秘过程。“沉迷”、“昏睡状态”等与“催眠”有关的词语都暗示着催眠是一种与清醒和睡眠都不同的独特意识状态。很多心理学家都在不同程度上同意这种观点。然而，尼古拉斯·斯潘诺斯（Nicholas Spanos，1942年~1994年）却持相反意见。他认为，催眠实际上只不过是一种提高了执行某些特定行为的动机状态，无须依靠“昏睡状态”或意识状态的改变就可以完全将其解释清楚。

最初的催眠术要追溯到18世纪中叶，那时候人们才第一次把精神疾病归因于心理而不是机体病变。弗朗兹·安东·麦斯麦（Franz Anton Mesmer，1733年~1815年），是帮助心理学从巫术领域中脱离出来的传奇人物之一。他认为“癔症”是体内一种流经全身的磁性体液失去平衡的结果。他在实验室举行特殊的聚会，音乐缓缓响起，灯光逐渐变暗，麦斯麦穿着邓布利多（Dumbledore）（系列丛书、电影《哈利·波特》中的人物）一样的衣服，从装有各种化学药品的瓶子中取出沾有药液的小棒触碰被痛苦折磨的病人。他相信这将把那些化学药品中他称为“动物磁性”的东西传递到病人那里，以使他们的病症得到缓解。有趣的是，历史上记载了很多用此方法治疗成功的个案（可能是安慰剂效应）。正是从麦斯麦那里，我们获得了“麦斯麦术”（mesmerize）这个名词，许多人认为这种治疗方法包含了很多与现代催眠术有关的技术。

纵观心理学的历史，特别是在心理治疗领域，催眠术（以希腊睡神Hypnos命名）占据着很重要的地位，而且它还是弗洛伊德精神分析技术的主要组成部分。有人认为催眠是一种改变了的心理状态，欧内斯特·希尔加德（Ernest Hilgard，1904年~2001年）则始终站在支持这一观点的研究者的前列（Hilgard，1978；Kihlstrom，1998）。他和其他一些人都认为催眠术包括了如下特征：对暗示的敏感性提高、不由自主地做出某种行为、记忆力提高、视觉表象急剧增加、意识分裂（能意识到某些经历，而对另一些却毫无意识）、痛觉缺失（对疼痛的感受性很低），等等。直到上20世纪70年代，人们关于“催眠可以让人产生正常状态下不可能有的想法、念头和行为”以及“催眠是一种不同的意识状态”的观念仍然是毋庸置疑的。

然而，科学家们必须经常用批评的眼光来看待研究现状，一旦看到

问题，就有可能推翻一些共识。社会心理学家尼古拉斯·斯潘诺斯（Nicholas Spanos）提出，希尔加德等人构建的有关催眠的主要假设应受到质疑。斯潘诺斯在文章中写道："主张催眠行为受控于特殊过程不仅是不必要的而且可能是一种误解……催眠行为基本上与其他社会性行为没什么区别，并且像其他社会性行为一样，催眠行为可以说是有策略或有目标的。"（P.200）换句话说，斯潘诺斯主张催眠行为实际上是一种为求达到特定结果的自主行为，被催眠者借助这种行为以达到所需要的结果。他进一步指出，如果这种行为可能由较高的动机水平引起，那么催眠就不涉及对意识状态的改变。

理论假设

斯潘诺斯推论，那些所谓催眠状态下的行为，实质上都在人正常的自主能力范围之内。他指出，一个人确定自己被催眠的唯一原因是：他们在催眠条件下的行为与自己期望在这种状态下会出现的行为相一致。在斯潘诺斯看来，催眠的过程是西方文化中一种具有多种含义的仪式。被试希望放弃对自己行为的控制，随着催眠过程的深入，他们开始相信他们的自主行动开始转化为自发的不随意活动。斯潘诺斯举了一个这样的例子：在催眠过程的初期，给被试一些随意活动的指导语，如"放松你的腿部肌肉"，但后来就变成一些不随意活动的暗示，如"你的腿感觉到沉重无力"。

在1982年发表论文之前，斯潘诺斯和他的同事、助手研究了近十年，力图证明常见的催眠效应可以通过并不神秘的简单方式来加以解释。

方法

本文并不介绍某个特定实验，而是总结了斯潘诺斯等人在1982年以前的很多研究成果。这些成果支持了斯潘诺斯等人的观点，反驳了希尔加德的论点（以及人们的普遍看法）——催眠是一种独特的意识状态。在此报告中，绝大多数研究成果分别来自于斯潘诺斯直接参与的16项研究，它们对催眠所产生的行为给出了另一种解释。因此，与前面关于梦境研究的论文一样，研究的结果和讨论将合并论述。

结果和讨论

斯潘诺斯声称，催眠术中有两个关键因素，致使人们相信催眠是一种"改变了的意识状态"：一是被试把自己的行为解释为是由别的什么东

西引起的，而不是他们自己引发的，这样就会使某些动作看起来是非自主行为；二是前面已谈到的观点，即催眠仪式让被试产生了一种期望，这种期望就促使被试以与期望相一致的方式表现某种行为。在本文中，斯潘诺斯将研究焦点集中在怎样对那些常被引用的催眠术观点提出质疑。

被催眠者的行为是非自主的

当被试被催眠时，主试会经常要他们做各种测试，以确定是否已进入了催眠状态。斯潘诺斯称这些测试经常使用一些特殊的方式，诱使被试确信自己有一些不同寻常的事情正在发生。催眠测试包括如下暗示："你的胳膊很沉重，你无法抬起它"；"你的双手正被某种力量拉到一起，而你无法把它们分开"；"你的胳膊会像铁棒一样僵直，无法弯曲"；或是"你的身体太沉，你无法站起来"。斯潘诺斯解释说，这些测试中的暗示包含了两个相关的要求：一个是要求被试去做某些动作，另一个是要求他们把这些动作解释为不随意行为。这些暗示对有些被试完全不起作用。斯潘诺斯称这些被试并不知道他们必须要随意地做一些动作才能引发暗示所指的行为，而不只是等待胳膊或身体自己开始运动。虽然另一些被试对暗示作出了反应，但他们发现这些行为是随意的。最后，也有一些被试满足了这两种要求，他们对暗示作出反应，并认为它们并不受他们控制。

斯潘诺斯提出，被试把自己的行为解释为随意的还是不随意的，这主要取决于暗示性指导语的措词方式。在另一项研究中，斯潘诺斯让两组被试通过诱导进入催眠状态。对其中的一组，主试给予多种行为暗示，例如，"你的胳膊很轻并正在抬起"。对另一组被试则直接指导他们做出同样动作，如"举起你的胳膊"。随后询问被试，他们的行为是否是随意的，受到暗示的那组被试与接受直接指导的那组被试相比，更倾向于认为自己的行为是不随意的。

读到这里，请你马上向前伸直你的左臂并保持几分钟。你会发现它开始渐渐变沉。这种变化不是由催眠引起的，而是重力的作用！所以，如果你被催眠，并受到暗示说你那伸直的手臂开始变沉，你就会很轻易地将它归因于某种不随意的力量（无论如何，你是自己想把胳膊放下来的）。但是，如果你受到的暗示是你的胳膊很轻并正在抬起的话，又会怎样呢？如果你抬起了胳膊，那你就很难把那个动作解释为不随意的，因为你无法忽视重力发出的矛盾反馈。斯潘诺斯检验了这种想法，并发现这样的解释是非常困难的。相信自己被催眠的被试，在把放下胳膊与抬起胳膊的动作相比较时，显然更倾向于认为放下胳膊的动作是不随意的。传统的催眠理论认为，催眠暗示中胳膊运动的方向没有受到任何影响，胳膊的动作总被认为是不随意的。

用暗示催眠被试，常常要求他们去想象特定情境以便他们产生主试想要的行为。如果你是一名被试，你可能会接受这样一个暗示："你的胳

膊很僵硬，不能弯曲。”为了强化这种暗示，主试还可能会加上“你的胳膊涂上了石膏一样沉”。斯潘诺斯相信，某些人比其他人更会被这些雕虫小技所影响，从而相信他们的反应（胳膊不能活动）是不随意的。他的理由是，如果你高度专心于此，你就无法注意到另外一些信息，而只有那些信息才会提醒你这些幻觉并不真实。你把胳膊上的石膏想象得越生动，精确到它的质地和硬度，以及它是如何被装到胳膊上去的，等等，你就越不会想到这只是你的想象力在起作用。如果你深深地专注于想象中发生的这些事情的细节，就更有可能倾向于相信僵直的手臂动作是不随意的，尽管事实并非如此。为了支持这一观点，斯潘诺斯发现，在要求被试评价自己对想象情景的专注程度时，对专注程度评价越高的人，就越容易认为自己的行为是不随意的。斯潘诺斯也注意到，对催眠很敏感的人与他（或她）沉醉于诸如读书、听音乐或做白日梦等活动中的倾向密切相关。因此，这些人更乐于接受催眠中的暗示。

被催眠的被试会产生某种期望

斯潘诺斯宣称，大多数相信催眠的人，也恰恰容易使他们产生典型的催眠行为。他进一步提出，用于诱发和研究催眠的方法能加强这种信念。他举了三个研究实例，这些例子都证明人们之所以在催眠状态下做出某种特定行为，是因为他们相信催眠就应该是那样，而不是因为催眠改变了他们的意识状态。

斯潘诺斯首先提到一项研究是给两组学生做关于催眠术的讲座。讲座中，除了告诉其中一组学生胳膊僵直的感觉是在催眠过程中不由自主出现的现象以外，其余内容完全相同。随后，两组学生都被催眠。在讲座中听到有关胳膊僵直的信息的那组学生，在没有给出任何指导语的情况下，就有人不由自主地表现出了这种行为。然而，在另一组被试中，无一人胳膊变得僵直。按照斯潘诺斯的观点，这就说明在催眠中，人是按自己对催眠的想象来行动的。

支持斯潘诺斯观点的第二项研究有以下发现：被试声称他们在催眠状态下产生的视觉表象要比未被催眠时产生的更为强烈、生动和真实。实质上，从这些研究的实施过程就能得到答案。主试要求被试想象他们正在某种背景或情景下完成特定动作，然后对相同被试进行催眠，要求他们再去想象相同或相似的情景（催眠和不催眠的实验次序可任选）。这些被试普遍报告在催眠条件下产生的表象更为强烈。然而，斯潘诺斯及其助手们发现，若使用两组被试，一组被催眠，另一组不催眠，其视觉表象的平均强度值近乎相等。为什么会出现这种差异呢？这一现象可以这样解释，即后一种方法使被试没有任何东西可以用来作对比。然而，如果我们在两种测试条件下使用同一被试，那我们才能够将两种条件下的感受进行比较，并依据一个来评价另一个。因此，既然被试几乎总是把催眠状态下的表象评价为更强烈，那么这是否说明催眠的确是一种“改变了的意识状态”呢？如果你问斯潘诺斯，他会说：“不！”按照他的观点，参加了两种条件测试的被试期望着催眠过程会带来更强烈的表

象，因此，他们是根据自己的期望作出评价的。

由斯潘诺斯援引的第三项研究对催眠的解释也许是最有趣的，即催眠可以使人们对疼痛变得不敏感（痛觉缺失效应）。通过“冷加压测试”可以在实验室中检测出被试的疼痛情况，但又不会对被试造成任何伤害。如果你是这项研究的被试，主试就会要求你把自己的胳膊浸入冰水中（零摄氏度），并坚持到你不能忍受为止。约经过10秒左右，胳膊的痛感明显加强，大多数被试会在1到2分钟内把手臂移开。希尔加德（1978）指出，在两种状态下接受痛觉缺失训练的被试报告，他们在催眠状态下感觉到的冷加压疼痛显著小于清醒状态。他对此的解释是，在催眠状态下，人可能会把疼痛从意识中分离出来。希尔加德强调，人的一部分意识体验到疼痛，但“健忘障碍”却将这种意识隐藏起来。

斯潘诺斯又一次对痛觉缺失的催眠性解释予以驳斥，并提供证据证明，在催眠时被试疼痛感的降低是他们动机和期望作用的结果。所有催眠研究使用的被试都是在受催眠影响的敏感性测量中获得较高分数的人。斯潘诺斯认为，这些人“非常投入，希望自己在催眠实验中成为一个好的被试”（P.208）。这些被试知道主试正在把清醒状态与催眠状态做比较，被试们想证明催眠的确有效。斯潘诺斯完成了一项含有冷加压疼痛的类似研究，但与前面的研究有一个很大的不同：主试告诉所有被试，他们将先使用清醒时的痛觉缺失技术（例如自我分心），然后告诉一部分被试，还将进行催眠状态下的疼痛减轻测试，但另一部分被试不被告知这一点。

图2—7是对斯丹姆斯潘诺斯研究发现的总结。当被试期待着催眠状态下的测试时，清醒状态下的测试完成后，他们会把痛觉缺失的效果评价得更低，正如斯潘诺斯所言，他们是为了给催眠状态下测试效果的提高“留有空间”。斯潘诺斯指出，这就证明了对疼痛不敏感的催眠行为可能是由于被试需要对情境要求作出反应，而不是自动产生的一种意识分离状态。

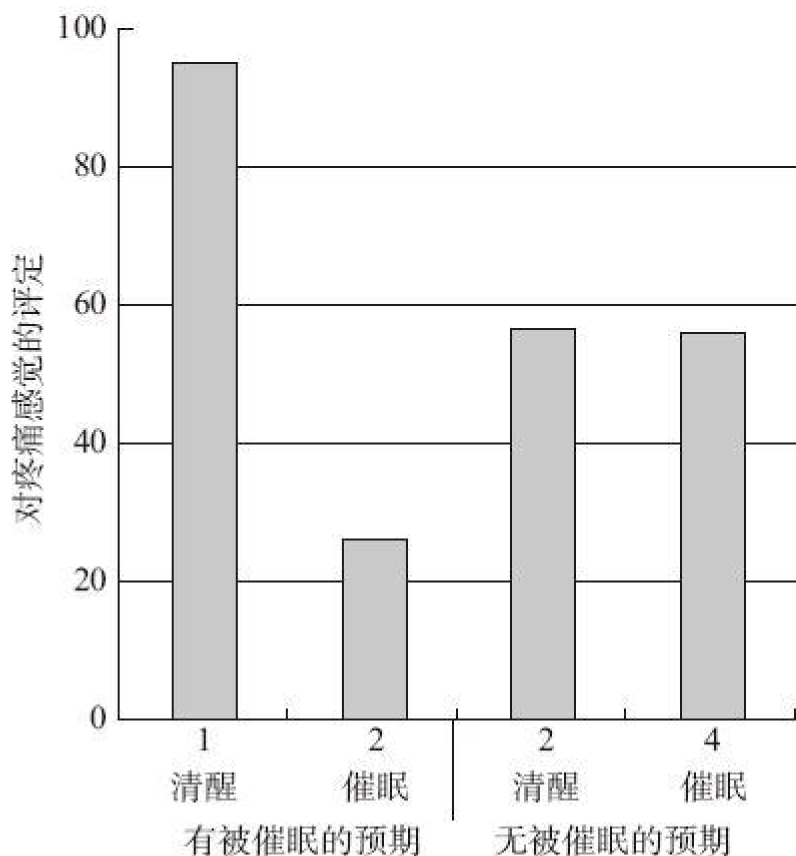


图2—7 清醒状态与催眠状态下痛觉缺失的比较：有预期与无预期

围绕着斯潘诺斯所报告的这些实验结果，最重要的疑问是：我们是否应该重新评价所谓的“催眠”现象。如果我们确认催眠不是像一般人和很多心理学家描述的那样，是一种有力的心理改变力量，那么它又将意味着什么呢？

研究发现的意义

评价斯潘诺斯的研究时，必须记住他的目的不是要证明催眠现象是不存在的，而是要证明催眠影响的行为是一种由较高的动机和目标驱动的社会性行为，而不是意识状态的改变或者催眠是一种独特的意识状态。有一点已经被大多数行为科学家所接受，即人不可能违背自己的愿望而被催眠。并且，在催眠状态下，被试不会做出他们认为是反社会的行为，也不可能表现出超人的力量或耐力。斯潘诺斯在本文中已证明，催眠现象的很多微妙内容都可以用并不神秘或更为直接的方式来解释，而不再以催眠后的意识状态来解释。

接受斯潘诺斯的“催眠并不存在”的观点意味着什么呢？对这个问题的回答是：“也许不意味着任何东西。”无论催眠的效果是由一种改变了的意识状态产生的，还是由提高动机水平产生的，均不会改变这样的事实，即催眠常是一种帮助人们改善生活中某些方面的有效方法。毫无疑问，催眠的作用已为人们广泛接受，其原因之一是，如果其他改善生活的方法都失败了，它就是人们最后一种解决其问题的途径——它的力量如此强大，以至于能使人迎着自身对这种变化的抵抗而进行改变。

有关催眠是否是一种“改变了的意识状态”的争论仍在继续。但无论催眠是什么，它毕竟不是大多数人乐于寻找的灵丹妙药。多项研究均表明，与其他治疗方法相比，催眠在帮助人们戒烟戒酒、提高记忆力或者减肥等方面并不是特别有效（研究回顾见Lazar & Dempster，1981）。

近期应用

如果有一个机会，你会愿意被催眠，还是拒绝被催眠？有人（Capafons et al.，2005）在斯潘诺斯研究的基础上，对那些对催眠持怀疑态度的人进行了一项研究。正如你将要看到的，这项研究结果的发现既能用来反驳斯潘诺斯的观点，也能用来支持他的观点。对于这些犹豫的被试，一旦他们确信同意参加催眠的实验，就会被分为三组：第一组被试几乎不提供任何在之后的催眠过程中会经历什么的信息（作为控制组）；第二组会提供关于随后催眠过程的“认知-行为”信息，即在随后的催眠过程中如何思考和如何活动的期望；第三组中为意识恍惚组，告知被试催眠是由于认知分离所造成的一种意识状态的改变（相对于正常的意识状态来说是一种独立的意识状态）。

在催眠开始的阶段，这些被试中哪些更有可能表现出回避反应呢？有些人可能认为，认知—行为组和意识恍惚组更有可能对催眠过程产生抵抗，催眠结果也会更加不明显（或者不能被催眠）。但是，令人惊奇的是，结果却与一些人的推测相反。认知-行为组和意识恍惚组都比控制组表现出了更高的接受催眠的暗示性。这些结果说明了什么呢？结果的意义也许并不明确。这可能印证了那句老话“知识就是力量”，体验到了一种更强的控制感，在催眠的过程中更能对自己进行“控制”。

另一项研究引用斯潘诺斯的催眠理论，对有些心理治疗师在咨询实践中的行为提出批评。这些治疗师常诱导来访者回忆出表面上“被压抑的”自己过去遭受性虐待的记忆（Lynn et al.，2003）。作者认为，催眠以及其他有争议的治疗技术可能会歪曲记忆，甚至创造出受虐待的记忆。实际上，这些记忆中的事情可能从来没有发生过，特别是在儿童早期的经历中（要了解更多内容，请阅读第4章Elizabeth Loftus关于记忆恢复的问题）。研究者指出，根据斯潘诺斯的研究，“成人对两岁或者更早之前事情的记忆，可能代表了早期事件的虚构、浓缩和构建，也代表当前所关心的内容以及所听到的关于早期事件的故事”（Lynn et al.，

2003，P.42）。换句话说，认为催眠技术可以使来访者提取出自己早期创伤性经历的准确记忆是一种误导，这种记忆会像在非催眠状态下的记忆一样，出现各种错误。作者认为，这可能导致对从未发生过的受虐事件错误的记忆和指控。斯潘诺斯在他1994年出版的书《多重身份和错误记忆：社会认知观》（Multiple Identities & False Memories: A Sociocognitive Perspective.）中阐述了他关于催眠技术中滥用潜意识的观点。

结论

显然，争论仍在继续。斯潘诺斯在1994年6月由于飞机失事而英年早逝，在此之前，他一直在继续他的研究（见McConkey & Sheehan，1995）。在他1988年的著作《催眠：认知行为观》（Hypnosis：The Cognitive-Behavioral Perspective）一书中，可以找到其早期催眠研究的总结。尼古拉斯·斯潘诺斯是一位成果丰富、令人尊敬的行为科学家，他的同事和那些受益于其研究的人们都将深深地怀念他（见Baker，1994，缅怀尼古拉斯·斯潘诺斯）。显然，人们会继承他的遗志以及他留下的宝贵遗产。他那关于催眠的研究改变了心理学，因为在将近二百年的时间内，对人类意识和行为这一方面的认识从未受到过任何实质性的挑战，而斯潘诺斯基于实验对此提出了一个全新的解释。

Baker, R.(1994).In memoriam: Nick Spanos.Skeptical Inquirer，18(5), 459.

Capafons, A., Cabañas, S., Alarcón, A., Espejo, B., Mendoza, M.E., Chaves, J.F., & Monje, A.(2005).Effects of different types of preparatory information on attitudes toward hypnosis.Contemporary Hypnosis, 22, 67-76.

Hilgard, E.(1978).Hypnosis and consciousness.Human Nature，1, 42-51.

Kihlstrom, J.F.(1998).Attributions, awareness, and dissociation: In memoriam Kenneth S.Bowers，1937-1996.American Journal of Clinical Hypnosis, 40(3)，194-205.

Lazar, B., & Dempster, C.(1981).Failures in hypnosis and hypnotherapy: A review.American Journal of Clinical Hypnosis, 24, 48-54.

Lynn, S., Loftus, E., Lilienfeld, S., & Lock, T.(2003).Memory recovery techniques in psychotherapy: Problems and pitfalls.Skeptical Inquirer, 27, 40-46.

McConkey, K., & Sheehan, P.(1995).Nicholas Spanos: Reflections with gratitude.Contemporary Hypnosis , 12, 36-38.

Spanos, N.(1994).Multiple identities & false memories: A sociocognitive perspective.Washington, DC: American Psychological Association.

Spanos, N., & Chaves, J.(1988).Hypnosis: Thecognitive-behavioral perspective.New York: Prometheus.

Stam, H.J.& Spanos, N.(1980).Experimental designs, expectancy effects, and hypnotic analgesia.Journal of Abnormal Psychology, 89, 751-762.

第三章 学习和条件反射

心理学在学习和条件反射方面已经拥有了一系列为数众多的优秀研究，这些研究关注的是动物和人的学习过程。心理学史上一些最著名的人物——不论他们是否属于行为科学领域，如巴甫洛夫、华生、斯金纳和班杜拉等——都在该领域中作出了最具影响力的贡献。尽管从心理学的这一分支中选取一些最有影响的研究绝非易事，但本章所选的内容几乎都可以在所有的入门心理学教材中找到，并且对这些科学家的巨大成就来说也极具代表性。

首先，要谈谈巴甫洛夫，我们将追溯到一百多年前，去重温他用狗、节拍器、唾液分泌装置所做的研究，以及他对条件反射的发现；其次是华生，以诸多贡献而闻名于世（或者说是声名狼藉？），最有名的大概就是他在20世纪早期用小艾尔伯特做的极具伦理挑战性的实验，该实验首次证明了人的情绪是环境的产物，而不仅仅是内在的过程；本章第三个研究讨论了斯金纳对鸽子迷信行为的著名解释和论证，以及人类是如何以同样的方式变得迷信的解释；最后，介绍了著名的“波比娃娃研究”，通过该实验，班杜拉指出儿童的攻击性行为是通过模仿成人的暴力行为而习得的。

研究9 不只限于分泌唾液的狗

Pavlov, I.P.(1927).Conditioned reflexes.London: Oxford University Press.

当你走进牙医诊所时，那里的消毒水气味是否曾令你感到牙疼？如果有，这可能是由于这种气味引发了你的某种联想，即该气味和以前看牙医的经历之间在你的大脑中形成的一种条件反射。当你听到美国国歌《星条旗永不落》在奥林匹克运动会上奏响时，你是否会感到心跳加速？大多数美国人都会这样。当你听到意大利国歌的时候，是否会发生同样的状况？除非你是在意大利长大的，否则你的心跳一般不会加快。这是因为你已经条件性地对一首国歌作出反应，而对其他国歌则不会。当你在某些人身边吹气球时，为什么他们会眯着眼且变得很紧张？很显然，这是因为他们已经把膨胀的气球和一些可怕的事情（如一声巨响）联系在了一起。上述行为只是无数由“经典条件反射”形成的人类行为中的几个例子。

关于学习的经典条件反射理论，是在大约一百年以前，由心理学史上最著名的人物之一——俄国的伊万·彼德维奇·巴甫洛夫（Ivan Petrovich Pavlov，1849年～1946年）提出并发展起来的。与本书中提到的大部分研究不同，巴甫洛夫的名字及其联想学习的基本观点在流行文化中被广泛接受（甚至在一首20世纪70年代滚石乐队的歌中也唱到“我像巴甫洛夫的狗一样流着口水”）。然而，人们大多并不清楚巴甫洛夫完成这一里程碑式发现的过程，也不理解其工作的真正意义。

虽然巴甫洛夫对心理学作出了重要贡献，但他并不是一名心理学家，而是研究消化过程的俄国著名生理学家。1904年，他因在消化方面的研究而获得诺贝尔奖，而那些使他的职业生涯和心理学史发生巨大改变的重要发现却始于偶然。值得注意的是，在19世纪末，心理学还是一门十分年轻的科学，而且很多人认为它不是真正的科学。因此，对巴甫洛夫来说，从更可靠、更令人信服的生理科学转向心理学是一种很冒险的激进转变。他在20世纪初曾写下了生理学家在转行去研究大脑和行为时所面临的两难境地：

> 从逻辑上讲，生理学家对生物行为的分析应该建立在更先进、更精确的科学，即物理和化学的基础上。但是如果要研究心理科学……我们将会把上层建筑建立在这门并不精确的学科上……事实上，心理学是否是一门自然科学，或者是否可以被看作一门科学，仍是一个值得探讨的问题（P.3）。

回顾巴甫洛夫的科学发现，他冒着这种风险并作出了激进性的转变。这对于心理科学的进步和对人类行为的理解来说，都是一种幸运。

巴甫洛夫的生理学研究主要是以狗为被试，研究唾液在消化中的作

用。他和他的助手把各种可食用或不可食用的东西放入狗的嘴里，以观察唾液分泌的比例和数量。为了科学地测量唾液量，他们在狗身上做了个小手术，即在狗的面颊上划一小口，插入一根导管，用来收集唾液。通过这一研究，巴甫洛夫得到了很多崭新而有趣的发现。例如，他发现当一只狗得到流质食物后，只有少量唾液分泌；相反，当研究者给狗喂干性食物时，就会有大量唾液产生。在这些条件下，分泌唾液被巴甫洛夫称为“反射”，它是一种对特定刺激自动发生的反应，不需要任何学习。对人来说，分泌唾液也是一种纯粹的反射。假如你在读这句话时，我要求你尽可能多地分泌唾液，你一定做不到。但是，如果你饿了，并发现面前有诱人的食物时，不管你愿意与否，你都会分泌唾液。

随着研究的继续，巴甫洛夫开始注意到一些完全出乎意料的现象。狗在食物放入嘴里之前，甚至在闻到食物的气味之前，就已经开始分泌唾液。一段时间以后，在没有任何消化刺激时，狗有时也会分泌唾液。这种分泌唾液的反射行为会随着狗在实验室中的某种经历而改变：“即使只给狗呈现喂食的容器，也足以引起它的完全消化反射（分泌唾液）；而且，狗看到喂食的人或者听到他的脚步声，也会产生这种反射”。（P.13）

这对巴甫洛夫来说，是面临抉择的重要关头。他观察到与消化似乎并无关系的刺激引起了消化反射，用纯生理学理论无法解释这种现象。问题的答案只能到心理学中去找。

理论假设

巴甫洛夫提出，在实验室中，狗从经验中学会了在某种信号出现之后，期待食物的出现。这些“信号刺激”并不会自然地引起唾液分泌，但狗却把它们和食物联系起来，并且作出唾液分泌的反应。因此，巴甫洛夫认为一定存在着两种类型的反射。

“无条件反射”是先天的、自动的，无需学习，并且对同一物种的所有成员而言都大体相同。食物进入口中时，人会分泌唾液；听到巨大的声音时，人会跳起来；灯光熄灭后，人的瞳孔会放大，等等，这些都是无条件反射的例子。条件反射则正好相反，它是通过经验或学习获得的，并在同一物种的不同成员间可能存在很大的差异。狗在听到脚步声时分泌唾液，或者当你闻到牙科诊室的气味时，你会感到牙疼等，这些都是条件反射。

无条件反射的形成是无条件刺激（UCS）引起无条件反应（UCR）。在巴甫洛夫的研究中，无条件刺激（UCS）是食物，无条件反应（UCR）是唾液分泌。条件反射的形成是条件刺激（CS），如脚步声，引起条件反应（CR），如唾液分泌。你可能注意到这两个例子中的反应都是唾液分泌，但若唾液分泌的原因是听到脚步声的话，便是一种

条件引起的反应，这是一种学习，不是狗的本能。

巴甫洛夫试图解决的问题是：条件反射既然不是先天性的，那么它究竟是怎样获得的呢？他认为当给狗喂食时，如果在狗所处的环境中经常出现一种特定刺激，那么在狗的大脑中，这种刺激将与食物建立联系，从而成为食物即将要到来的信号。在食物与刺激建立联系之前，这种环境刺激不产生任何重要反应。换言之，对于狗来说，它是一种中性刺激（NS）。当狗一旦进入实验室时，虽然实验人员的脚步声可能使狗产生一种好奇反应（巴甫洛夫称它为“是什么”反应），但是它并不能引起狗的唾液分泌，此时的脚步声为中性刺激。然而一段时间后，因为每天狗被喂食前都听到同样的脚步声，它们就开始把这种声音与食物联系起来。根据这种理论，最终，仅仅出现脚步声也能引起狗的唾液分泌。所以，依照巴甫洛夫的观点，表3—1列出了中性刺激转变成条件刺激的过程。

表3—1 中性刺激转变成条件刺激的过程

第一步			UCS（食物）	→	UCR（唾液分泌）
第二步	NS（脚步声）	+	UCS（食物）	→	UCR（唾液分泌）
第三步	多次重复第二步				
第四步			CS（脚步声）	→	CR（唾液分泌）

这样巴甫洛夫就有了一套理论来解释他所观察到的现象，于是他便开始着手进行一系列实验，来证明这种理论的正确性。我们通常认为，在巴甫洛夫实验中，狗是听到铃声而分泌唾液的，这在后续的实验中被证明是正确的。但是正如你将看到的，在早期实验中，他使用的是节拍器。

方法和结果

巴甫洛夫得到了“一位热心而有公益精神的莫斯科商人”的资助，在彼得格勒（后来成为列宁格勒，现在又改成了它最早的名字——圣彼得堡）的医学实验研究中心建立了一个特殊的实验室。这个隔音实验室能够使被试与实验者及实验过程中的所有额外刺激完全隔离开。因此，在那里可以使用一种特定的刺激，并且在记录反应时，可排除实验者和动物直接接触而产生的影响。

在建立了这种严格控制的实验环境之后，实验程序就相当简单了。巴甫洛夫选择食物作为无条件刺激。正如前面所提到的，食物将引起唾液分泌这一无条件反应。然后，巴甫洛夫需要寻找一种对狗来说是与食物毫不相干的中性刺激。为此，他选择了节拍器的声音。在几次建立条件反射的试验中，狗听到节拍器的滴答声后，就会立即得到食物。“一种中性的刺激被强加在先天的消化反射行为之上。我们观察到，这种刺激

的组合在重复了几次以后，节拍器的声音就具有了能刺激唾液分泌的属性”（P.26）。换句话说，对于分泌唾液这种条件反应来说，节拍器的声音已经成为它的一种条件刺激。

巴甫洛夫和他的助手通过使用不同的无条件刺激和中性刺激来详细说明这种初步发现。比如，在把一种柠檬汁之类的溶液倒入狗的嘴里之前（无条件刺激），向狗呈现香草的气味（中性刺激）。果汁无疑会引起狗的大量唾液分泌（无条件反应）。这种组合大约重复呈现20次后，即使只呈现香草的气味，狗也会分泌唾液。在一个视觉实验中，主试在给狗喂食之前先旋转一个物体，这种组合在重复5次以后，仅出现旋转的物体（条件刺激）就能引起狗的唾液分泌（条件反应）。

巴甫洛夫工作的重要性和应用性不仅限于对狗分泌唾液的研究。他的经典条件反射理论解释了人类行为的很多方面，并有助于推动心理学成为一门真正的科学。

研究发现的意义

经典条件反射理论（也称巴甫洛夫条件反射学说）被人们普遍接受，并且这一理论自巴甫洛夫的研究之后，一直未曾改变。它可以解释和说明许多人类行为，包括：恐惧症从何而来；你为何不喜欢某种食物；什么是你情绪的来源；广告是如何发生作用的；为何在面试或考试前你会感到焦虑；是什么引起你的性欲，等等。下面将讨论几项有关这一理论的研究。

经典条件反射关注反射行为：这些行为不为意识所控制。任何一种中性刺激都可用于形成条件反射。通过建立经典条件反射，你会在听到门铃声时眨左眼，看到蓝色闪光时心跳加快，或是吃草莓时产生性欲冲动。门铃、蓝光和草莓都是与条件反应无关的中性刺激，直到它们与无条件刺激配对联系起来时才成为条件刺激。这些条件反应即眨眼（如对着眼睛吹气）、心率加快（如突然一声巨响）和性冲动（如浪漫的拥抱）。

为了直观地体验经典条件反射的过程，你可以在自己身上进行以下这个实验。你所需要的只是一只铃铛、一面镜子和一间关灯后一片漆黑的房间，这个房间就是你的临时实验室。你的瞳孔会由于灯光亮度的变化而产生放大和缩小的反应。你无法随意地控制瞳孔变化，而且你也不必去学怎样控制它。如果我对你说“现在请你放大瞳孔”，你是无法做到的。然而，当你走进漆黑的剧场，你的瞳孔会立即扩大。因此，光线的减弱对瞳孔放大来说就是一种无条件刺激，能产生无条件反应。在临时实验室里按响铃，随即关上灯，在黑暗中等待大约15秒后再打开灯，再等15秒后重复这一过程：响铃……关灯……等待15秒钟……开灯……将中性刺激（铃声）和无条件刺激（黑暗）构成的组合重复10到20次，确

保铃声只在突然的黑暗前出现。现在，打开灯，靠近镜子仔细观察自己的眼睛，铃响后你将看到即使没有灯光的变化，你的瞳孔也会有轻微的放大！于是，铃声变成了条件刺激，瞳孔放大变成了条件反应。

相关研究和近期应用

本书中还有另外两个实验应用了巴甫洛夫的经典条件反射理论。在下一篇文章中，约翰·华生运用了与巴甫洛夫训练狗分泌唾液的同样原理，使11个月大的小艾尔伯特条件反射性地对白鼠（以及其他有毛的动物）产生恐惧。通过这种方法，华生证明了情绪（如恐惧）产生的过程。在此之后，约瑟夫·沃尔普（Joseph Wolpe，参见第9章中的研究34心理治疗）应用经典条件反射的概念，找到了一种治疗强烈恐惧（恐惧症）的方法。他的研究依据的理念是，只有打破条件刺激和非条件刺激之间的联系，才能减少恐惧反应。

关于经典条件反射与恐惧这一方面的研究一直持续到现在。例如，一些研究发现，如果父母对某种事物感到恐惧，比如害怕蛇和蜘蛛，儿童无需直接接触这种事物，只需通过从父母那里得到的“替代性”条件反射，就能产生对同样事物的恐惧（Fredrikson，Annas，& Wik，1997）。在心理学和医学文献中，应用巴甫洛夫理论的例子不计其数，这里无法将它们一一罗列。因此，接下来我们将选取几项著名的研究成果加以讨论。

有一个普遍困扰全世界牧场主的问题，那就是狼和土狼等食肉动物经常会捕食他们的家畜。20世纪70年代初，人们进行了一项无需杀死食肉动物，而是运用巴甫洛夫条件反射技术来解决狼和土狼吃羊的研究（参见Gustafson等人，1974）。首先，给狼和土狼吃一些含有少量氯化锂的羊肉（非条件刺激），这种化学物质摄入体内后，可使动物感到不适。这些动物吃了羊肉以后，便感到头晕、恶心，并开始呕吐（非条件反应）。待痊愈之后，把这些饥饿的食肉动物与活羊放入同一围栏中。狼和土狼起初攻击羊群（条件刺激），但它们一闻到猎物身上的气味，便停止攻击，并尽可能地远离羊群。当围栏的门被打开时，这些狼和土狼竟然迅速地逃了出去！根据这项研究和其他相关的研究，现在牧场里普遍采用这个方法，即运用经典条件反射理论就可使狼和土狼远离牧群。

应用经典条件反射理论的一个潜在的重要研究领域是行为医学。有证据显示，应用巴甫洛夫原理可以改变免疫系统的活动。阿德和库恩（Ader & Cohen，1985）给白鼠提供一种含有糖精的水（白鼠喜欢这种水）。随后，他们将这种糖水与一种可减弱老鼠免疫系统能力的药物注射建立联系。然后，再给这些建立了条件反射的白鼠提供这种糖水，这时尽管不给它们注射药物，白鼠也会表现出免疫力抑制的征兆，即一种免疫力减弱的反应。当前人们正在着手研究反向过程是否也能行得

通，即是否可以通过建立经典性条件反射，形成提高免疫力的行为反应（主要涉及心理学中的心理神经免疫学分支）。总的来说，研究显示，利用经典性条件反射来提高人类的免疫系统反应的前景是非常光明的（Miller & Cohen，2001）。可以设想一下，在将来的某一天，我们只需通过把人置于非药物条件刺激中，就可以提高人对疾病的抵抗力。比如，假如你着凉了或得了流感，便可以播放“能提高自己免疫系统反应的音乐”——特定的经典条件刺激。随着音乐响起，你的抵抗力会对音乐刺激作出条件反应，即免疫力得到增强，从而战胜疾病。

巴甫洛夫的发现对当今心理学研究仍具有巨大影响，这一点可以从以下方面得到证实。自2000年以来，已有一千多篇科学研究引用了巴甫洛夫的理论。最近一项尤其引人注目的研究发现，个体在建立和消除条件反射时的心理状态对治疗经典条件性的非理性恐惧（恐惧症）会产生影响（Mystkowski等，2003）。研究者运用脱敏技术对害怕蜘蛛的被试进行治疗，其中一些被试在摄取咖啡因后接受治疗，而另一些则在摄取安慰剂后接受治疗。一周后，所有的被试接受重测——一些被试摄取咖啡因，而另一些被试摄取安慰剂。那些在治疗阶段摄取安慰剂而在后测中服用咖啡因的被试，以及那些在治疗阶段服用咖啡因而在后测中服用安慰剂的被试，均再次表现出恐惧反应。换言之，刺激特性的改变，减小了消除症状的效果。然而，对于那些在治疗阶段和后测阶段中都处于相同药物条件下（咖啡因或安慰剂）的被试来说，他们对蜘蛛的恐惧反应则持续降低。这项研究意味着，即使一个经典条件反射行为已经消失，但当这种条件刺激在一个新的、不同的情景中出现时，该条件反应会再次出现。

结论

通过这些实例，可以清楚地看到巴甫洛夫的影响在心理学界是多么广泛。很少有科学家在某一学科有如此大的影响，特别是对于心理学家而言。经典条件反射理论是支撑现代心理学发展的基础理论之一。如果没有巴甫洛夫的贡献，在经过数十年后，其他行为科学家也许还是会发现这些原理中的绝大部分。但如果巴甫洛夫没有下定决心冒险地改变了职业，没有进入19世纪那未经验证且饱受质疑的未知心理学领域，那么如此衔接密切、完善而又十分明晰的条件反射理论就不可能诞生。

Ader, R., & Cohen, N.(1985).CNS-immune system interactions: phenomena.Behavioral and Brain Sciences, 8, 379-394.

Fredrikson, M., Annas, P., & Wik, G.(1997).Parental history, aversive exposure, and the development of snake and spider phobias in women.Behavior Research and Therapy, 35(1), 23-28.

Gustafson, C.R., Garcia, J., Hawkins, W., & Rusiniak , K.

(1974).Coyote predation control by aversive conditioning.Science , 184, 581-583.

Miller, G., & Cohen, S.(2001).Psychological interventions and the immune system: A meta-analytic review and critique.Health Psychology, 20, 47—63.

Myskowski, J., Mineka, S., Vernon, L., & Zinbarg, R (2003).Changes in caffeine states enhance return of fear in spider phobia.Journal of Consulting and Clinical Psychology, 71, 243-250

研究10 情绪化的小艾尔伯特

Watson, J.B., & Rayner, R.(1920).Conditioned emotional responses.Journal of Experimental Psychology, 3，1-14.

你想过你的情绪反应是从哪儿来的吗？如果想过，那你也不是唯一一位这样的人。情绪起源的问题在整个心理学史上一直令许多行为科学家着迷，这在本书中就可以得到证实。本书有四项研究是与情绪反应直接相关的（第5章，Harlow，1958；第6章，Ekman & Friesen，1971；第8章，Seligman & Meier，1967；第9章，Wolpe，1961）。这项由华生和雷纳（Rayner）所做的条件性情绪反应实验，在近一个世纪前公之于世时，便产生了极大的影响力，并且这种影响力一直延续至今。当你翻开任何一本基础心理学教材或有关行为与学习的书籍时，会发现它们几乎都包含了这项研究。

这项研究的历史重要性不仅要归功于研究结果本身，而且它还开拓了心理学的新领域。如果我们回溯到20世纪初，感受一下那时心理学的状况，便会发现当时的心理学几乎完全被西格蒙德·弗洛伊德的研究所统治（参见第8章安娜·弗洛伊德的相关部分）。弗洛伊德精神分析理论关于人类行为的基本观点是：行为的动力来源于无意识的本能和童年早期被压抑的内心冲突。简单地说，就是行为和特殊的情绪是由内在的生理和本能过程产生的。

20世纪20年代，心理学界开始了一场新的运动，以巴甫洛夫和华生为代表的行为主义出现了。行为主义者的观点与精神分析学派的观点完全相反，他们提出，行为是通过外在的不同环境和情境刺激而产生的。因此，华生指出，情绪反应是我们对环境中某种特定刺激的条件反射。换句话说，人的情绪反应是习得的。华生相信所有人类行为都是学习和条件反射的产物，正如他在1913年的著名研究报告中所宣称的：

> 给我一打健全的婴儿，让他们在我设定的特定环境中成长，我可以保证，随机选出的任何一名婴儿，我都可以把他训练成为我所选定的任何类型的特殊人物，如医生、律师、艺术家、商界领袖或乞丐和小偷。

在那个时代，这是一种极具革命性的观点。大多数心理学家以及公众普遍不愿接受这些新观点。对情绪反应来说尤为如此，因为情绪反应似乎是产生于内部的，所以华生试图用实验证明特定情绪可以通过建立条件反射来产生，而不是由内在作用引起。

理论假设

华生提出，假设一种刺激使你自发地产生某种特定的情绪反应（如恐惧），并且这种刺激每次重复时都伴随着其他事物，如一只白鼠，那么白鼠就会在你的大脑中与恐惧建立联系。换言之，你最终会条件反射性地害怕白鼠（这种观点体现了巴浦洛夫的经典条件反射）。华生强调，人类并不是天生害怕白鼠，这种害怕是通过条件反射习得的。这就是他最著名的实验的理论基础，该实验的被试名叫小艾尔伯特。

方法和结果

被试艾尔伯特是一名9个月大的孤儿，从出生起就一直待在医院里。研究人员和医护人员认为他在心理和生理上都很健康。为了解艾尔伯特是否天生害怕某种特定刺激，研究者给他呈现白鼠、猴子、狗、有头发和没有头发的面具以及白色羊绒绵。研究者密切观察艾尔伯特对这些刺激的反应。艾尔伯特对许多动物和物体都感兴趣，愿意接近它们，并不时地触摸它们，但从未表现出丝毫的恐惧。因为这些东西不会引起恐惧，所以可将它们看作是中性刺激。

实验的下一步是要确定艾尔伯特对巨大的声音是否会产生恐惧反应。这并不困难，因为我们所有人，特别是婴儿，都会对突然出现的巨大声音产生恐惧反应。由于这种反应是无需学习就会发生的，所以巨大的声音被看作是“无条件刺激”。在本实验中，研究者在艾尔伯特身后用锤子敲一根4英尺长的铁棒。这种声音使他受到惊吓，并开始哭泣。

现在我们就可以检验艾尔伯特的恐惧情绪是否是条件反射的结果了。真正的条件反射检测是在艾尔伯特11个月大时进行的。研究者曾对通过实验引起儿童恐惧反应的做法感到犹豫，但最后他们决定继续进行（当时持有的伦理逻辑现在回想起来的确值得商榷，本章稍后将对此问题及该研究涉及的所有伦理问题进行讨论）。

实验开始后，华生和他的研究生助手罗莎莉·雷纳（Rosalie Rayner）向艾尔伯特呈现白鼠。一开始，艾尔伯特对白鼠很感兴趣，并试图触摸它。在他正要伸手时，铁棒被突然敲响，突如其来的响声使艾尔伯特感到十分惊恐，这一过程重复了3次。一周以后，重复同样的过程。在白鼠与声音同时呈现共7次后，不再出现声音，只向艾尔伯特呈现白鼠时。你可能已经猜到，艾尔伯特对白鼠产生了极度恐惧。他开始号啕大哭，转身背对白鼠，向远离白鼠的方向爬得飞快，研究者不得不冲过去抓住他，以免他从桌子的边缘掉下来。对于一种物体从没有恐惧到产生恐惧只用了短短的1周时间。

随后，研究者想探讨这种习得的恐惧是否会迁移到其他物体上。在心理学术语中，这种迁移叫作“泛化”。如果艾尔伯特对其他相似刺激物产生恐惧，那么这种习得的行为就已经泛化了。又过了一周，对艾尔伯特再次测试发现，他仍旧对白鼠产生恐惧。随后研究者欲检测这种恐惧

是否泛化，他们给艾尔伯特呈现一种与白鼠相似的东西（白兔）。研究者描述为：

消极反应立即出现，艾尔伯特尽可能地远离动物，低声抽泣，然后大哭起来。我们让他触摸兔子时，他却把脸埋在垫子里，然后用四肢将自己支撑起来，边哭边爬走了（P.6）。

这里要提醒读者的是，在这种条件反射建立前，艾尔伯特并不怕兔子，并且也没有建立对兔子的特定恐惧性条件反应。

在同一天内，研究者依次给小艾尔伯特呈现狗、白色毛皮大衣、一包棉花和华生头上的灰白头发。他对所有这些东西都感到恐惧。令这项研究声名远扬而又声名狼藉之处，就是华生把一个圣诞老人的面具呈现给艾尔伯特，这也是最著名的泛化实验之一。小艾尔伯特会是什么反应呢？没错，是恐惧！五天后，再次对艾尔伯特进行测试。表3—2中列出了这一天呈现的物品顺序。

关于条件性的情绪反应，华生想考察的另一方面是：习得的情绪是否会从一种情境迁移到另一种情境中。如果艾尔伯特对这些动物和物品的恐惧反应只发生在实验室而不发生在别的地方，那么其研究成果的价值将大大减弱。为了验证这一点，在进行表3—2中的测试的同一天，研究者将艾尔伯特带到一个完全不同的房间，那里灯光更明亮，而且人更多。在这种新环境中，很明显艾尔伯特仍对白鼠和兔子感到恐惧，只是不像以前那么强烈。

表3—2 测试第四天对艾尔伯特呈现刺激的顺序

呈现刺激	观察到的反应
1. 积木	和平常一样玩积木
2. 白鼠	害怕、后退（没有哭）
3. 白鼠＋噪音	害怕并哭泣
4. 白鼠	害怕并哭泣
5. 白鼠	害怕、哭泣并离开
6. 兔子	害怕，但不像前面表现得那么强烈
7. 积木	像平常一样玩积木
8. 兔子	与 6 相同
9. 兔子	与 6 相同
10. 兔子	有点儿害怕，但还想触摸它
11. 狗	害怕，回避
12. 狗＋噪音	害怕并离开
13. 积木	像平常一样玩积木

华生和雷纳想要做的最后一个实验是观察艾尔伯特新习得的情绪反应是否会持续一段时间。但不久，艾尔伯特就被人收养，并要离开医院，因此所有测试在31天后终止。31天后，再给艾尔伯特呈现圣诞老人的面具、白色皮毛大衣、白鼠、白兔和狗，他仍然对这些东西感到十分

恐惧。

华生和他的同事还计划对艾尔伯特重建条件反射（recondition），以消除他的这些恐惧反应。然而，小艾尔伯特在完成最后一个实验后就离开了医院，正如大家所知道的，矫正实验没能进行。

讨论和研究发现的意义

华生在这项研究及他的所有工作中有两个基本目标：（a）证明人类所有的行为源于学习和条件反射；（b）证明弗洛伊德的心理学理论——即我们的行为来自无意识——是错误的。虽然这项实验在方法上存在缺陷，并且严重违反了伦理道德（我们将在下一页讨论），但在很大程度令人信服地说明了情绪行为可以通过简单的“刺激—反应”手段而成为条件反应。这一发现对开创心理学的另一个主要学派——行为主义功不可没。在这里，某些像情绪一样复杂、个性化、为人类所共有的东西也服从于条件反射原理，就像巴甫洛夫发现的狗能学会听到响铃后分泌唾液一样。

对这项研究的一个合乎逻辑的推论是其他情绪——如愤怒、愉快、伤心、惊讶或厌恶等都可用同样的方式习得。换句话说，当你听到一首老歌时，会感到伤感；求职面试或公众演讲时，会紧张；春天到来时，感到愉快；听到牙医手中工具发出的声音时，感到害怕。这些现象产生的原因是，在你头脑中这些刺激已与特定的情绪建立了条件反射联系。其他一些情绪的更极端反应，如恐惧症和性恋物癖，也可以通过相似的条件反射序列来形成。

华生很快指出，他的研究成果与弗洛伊德及其后继者的精神分析理论相比，能更直接更简洁地解释人类的行为。就像华生和雷纳在论文中提到的，一个弗洛伊德主义者会把吸吮拇指当作追求快乐的本能表现。然而，艾尔伯特会在他感到害怕时吸吮拇指。拇指一放到嘴里，他的害怕就能缓解。因此，华生把吸吮拇指解释为一种阻碍恐惧产生的条件反射。

在这篇论文中，华生和雷纳对弗洛伊德学派如何解释艾尔伯特将来对白色皮衣的恐惧进行了预测。他们称，弗洛伊德学派的分析家可能通过梦来进行分析。这个梦显示艾尔伯特在3岁时想玩妈妈的假发，却遭到妈妈的严厉斥责（P.14）。华生之所以用艾尔伯特的例子主要是想说明，成年人的情绪困扰不能总归因于弗洛伊德学派所主张的少年时代的性创伤。

问题和批评

读到这里，你可能已经对实验者对这名无辜儿童的做法感到担心，甚至是气愤。这项研究明显违反了人类行为研究中现行的伦理学标准。现在任何一个以人为被试的研究机构都不会赞成这项研究。然而，在一个世纪以前，这种伦理标准并未正式出现。在早期心理学文献中，经常可以看到实验方法上存在违背伦理的情况。应该指出的是，华生和他的同事并不是惨无人道之人，况且他们面对的是一个未经探索的领域。他们承认在进行条件反射程序前的确犹豫了相当长的一段时间，但最终他们确定这是无可非议的，因为即便不做这个实验，在艾尔伯特离开医院的庇护后，一些类似的恐惧也会自然出现在这孩子身上。即便如此，撇开实验结果的重要性不谈，单说让一名孩子惊惧到如此地步，难道是恰当的吗？如今绝大多数的行为科学家都会对此予以否定。

在这项研究中，另一个与伦理学有关的重要问题是，艾尔伯特在离开研究环境时并没有消除恐惧。华生和雷纳在他们的文章中承认这种引起情绪的条件反射可能伴随人的一生。倘若果真如此，那么从伦理学角度上讲，很难为研究者任由这种恐惧伴随艾尔伯特成长的错误行为找到辩解的理由（天知道还会有多少类似于这种刺激的东西会引起艾尔伯特的恐惧）。

与此相关，一些研究者还批评了华生关于这些恐惧会无限持续下去的假设（Harris，1979）。另一些人则怀疑华生对艾尔伯特施行条件反射的有效性（Samelson，1980）。很多研究证明，由于其他经历或仅仅是时间的推移都可使条件反射行为消失。可以想象一下，假如在艾尔伯特5岁时，有人把一只宠物兔子作为生日礼物送给他，一开始他会感到恐惧（这无疑会使他的养父母感到困惑），但在与兔子的相处中，他发现并没有可怕的事情发生（如巨大的声响），于是他对兔子的恐惧很可能会慢慢减弱，直到对兔子不再产生任何恐惧反应。这是学习心理学中一个十分明确的过程——“消退”。人一生中不断地经历着学习与消退、建立与解除条件反射的过程，“消退”是其中的一个常规部分，但对艾尔伯特来说，却并非如此。

近期，一个由研究生组成的研究团队耗时7年终于揭示了这个名为“艾尔伯特”的孩子的真实身份，这是一个保守了数十年的秘密。艾尔伯特已被确认就是道格拉斯·梅利特（Douglas Merritte），是阿维利亚·梅利特（Arvilla Merritte）的儿子。华生研究期间，阿维利亚在约翰·霍普金斯大学医院工作，是“哈丽特·莱恩之家”的一名护士。她同意让自己的儿子作为被试，并得到了1美元的报酬（见DeAngelis，2010）。在这些新的发现中，最令人悲伤的是道格拉斯在6岁时死于脑积水（大脑内部及其周围液体积聚）。现在针对这种疾病有许多疗法，但在20世纪20年代，这种病通常是致命的。因此，艾尔伯特是否克服了华生对他建立的恐惧仍不可获知。此外，还需提到的是，在这项研究开展期间，华生由于与他的研究生助理罗莎莉·雷纳发生婚外情而被约翰·霍普金斯大学解雇。

近期应用

尽管华生1920年发表的文章存在伦理缺陷，却仍一直被许多研究所引用。这些研究涉及范围很广，从子女养育到心理治疗。其中一项研究对婴儿情绪表达中的面部表情进行了考察（Sullivan & Lewis，2003）。我们知道，人类的面部表情与相应的特定情绪是一致的（见第6章艾克曼的研究）。但是，这项研究被进一步扩展，考察了婴儿的表情是如何发展的，以及在他们年龄非常小的时候，这些不同表情所表达的含义。对婴儿面部表情含义的理解，有助于父母更好地和婴儿交流并照顾他们。作者指出，他们的研究目标是“向执业医生提供基本信息，以帮助他们及父母更好地识别所照料的婴幼儿的表情信号（P.120）。”这项研究对华生的研究成果的运用让人感到某种欣慰，因为华生对小艾尔伯特采用的颇有争议的研究方法最终发挥了作用，使我们能够更加敏感地觉察到婴儿的情感和需要。

正如前面曾提到过的，一种情绪，譬如害怕，它的极端形式能产生严重的消极结果，即恐惧症。很多心理学家认为，恐惧症是通过条件反射形成的，类似于小艾尔伯特对有毛动物的恐惧的形成过程（参见本书第9章Wolpe治疗恐惧症的研究）。很多关注恐惧症产生原因和治疗方法的研究都纳入了华生的研究，其中有一项研究从先天—后天的视角探讨了恐惧症，并得出引人注目的结果。华生的研究，显而易见是完全基于环境（或后天）决定论的观点，大多数人也会认为恐惧症是习得的。

然而，肯德尔、克尔柯沃斯基和波瑞斯可特（Kendler，Karkouski & Prescott，1999）却提供了相反的有力证据：恐惧症的发展可能包括大量遗传因素的作用。研究者以1700多名女性双胞胎为对象，研究了她们的恐惧症和莫名的恐惧（参见本书第1章中鲍查德对双胞胎的研究）。他们发现恐惧症中有很大一部分是由于遗传因素造成的。作者的结论是：虽然恐惧症可以通过个体的环境经验形成，但在恐惧症的形成中，家庭的生物学影响要比其环境性影响更重要。设想一下：天生具有恐惧！这种观点与华生的理论尖锐对立，可以说，又给贯穿心理学和行为科学领域的先天—后天之争点了一把火。

DeAngelis, T.(2010).Little Albert regains his identity.Monitor on Psychology, 41，10.

Harris, B.(1979).What ever happened to little Albert? American Psychologist, 34，151-160.

Kendler, K., Karkowski, L., & Prescott, C.(1999).Fears and phobias: reliability and habitability.Psychological Medicine, 29(3), 539-553.

Samelson, F.(1980).Watson’s little Albert, Cyril Burt’s twins, and the need for a critical science.American Psychologist, 35, 619-625.

Sullivan, M., & Lewis, M.(2003).Emotional expressions of young infants and children: A practitioner’s primer.Infants and Young Children，16，120-142.

Watson, J.B.(1913).Psychology as the behaviorist views it.Psychological Review, 20，158-177.

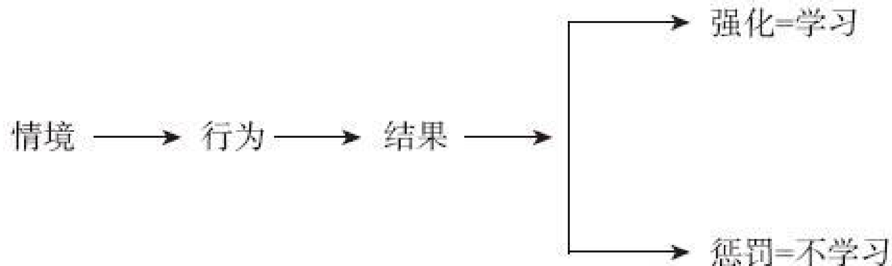
研究11 敲敲木头

Skinner, B.F.(1948).Superstition in the pigeon.Journal of Experimental Psychology, 38，168-172.

B.F.斯金纳（1904年~1990年）是影响深远、极负盛名的心理学家，他一生硕果累累。下面我们将对他的一项研究进行讨论。然而，该选择哪项研究来代表他的工作却是个难题，用短短的一篇文章来呈现斯金纳对心理学研究的历史贡献是远远不够的。毕竟，斯金纳被誉为激进行为主义之父，是著名的（或声名狼藉的）“斯金纳箱”的发明者，他还出版了20多本著作，发表了数百篇科学论文。我们将要探讨的这篇文章的题目是《鸽子的迷信行为》，之所以从他众多著作中选出这篇题目听起来很幽默的文章，是因为通过这个有趣的例子，不仅使我们清楚地认识斯金纳的基本理论，了解他研究行为的方式和方法，而且还提供了斯金纳对我们所熟知的“迷信”现象的解释。

斯金纳被称为激进行为主义者，因为他认为所有行为，包括公开的（或外部的）行为以及隐秘的（或内部的）行为（例如感情和思想），从根本上来说都是从诱发行为的情境与直接的行为结果间的联系中习得的，并且受这种关系所控制。尽管斯金纳认为隐秘的行为很难研究，但他也承认对这些行为我们都有自己的主观体验。然而，他并不把诸如思想和情绪等内部事件看成是产生行为的原因，而是将其看作是一种环境与行为相互作用的结果——这正是斯金纳想要解释的内容（对“激进的行为主义”的详细解释见Schneider & Morris，1987）。

简而言之，斯金纳的理论就是，在某一特定的情况下，你的行为都很可能会伴随着某种结果，比如得到赞扬、报酬或解决问题后的满足感，这些结果使得你在今后遇到类似的情况时，很可能会重复这一行为。这些结果被称为“强化物”。如果你的行为伴随着另一种结果，比如疼痛或尴尬，那么今后在相似的情况下，你很少会再重复这一行为，这些结果被称为“惩罚”。行为与环境相互作用的结果被分别称为“强化”和“惩罚”（Edward K.Morris的个人通信，1987年9月）。在斯金纳提出的“操作性条件反射”中，“强化”和“惩罚”是两个最基本的心理过程。这一过程图示如下：



通过这一概念，斯金纳也能解释习得的行为是怎样消退或彻底消失的。一种行为被强化后，强化物若不再出现，那么这一行为再发生的可能性将慢慢减弱直到完全消失，这种抑制过程被称作“消退”。

这些内容对你来讲，其实并不生疏。我们常用这种程序来训练宠物。你让狗坐下，它坐下了，你便奖励它一块美食。一段时间后，当你让狗坐下时，即使不及时给予食物奖赏，它也会坐下。这里你运用了操作性条件反射的原理。这是一种对所有动物都非常有效的学习方式，甚至可以让老狗学新技能，当然也包括猫！如果你想要宠物不再做某种动作，你要做的事就是把强化物挪走，随后这一行为便将停止。例如，如果你的狗蹲在餐桌旁乞食，这必然有原因（不管你怎么想，狗并不是生来就会在桌边乞食的）。这一行为是你通过奖赏的强化，让狗形成的条件反射。如果你想使这种行为彻底消失，就必须完全停止强化。最后，狗会停止乞食。顺便说一句，在消退期间，如果一个家庭成员偶尔偷偷地给乞食的狗一些食物，消退行为将永远不会发生，但是狗会花费更多的精力去接近这位家庭成员。

除有关学习的这些基本知识以外，斯金纳还认为人类的所有活动也以完全相同的方式产生并维持。只不过对于人类而言，识别行为及结果并不太容易。斯金纳的著名观点是：如果一个人的行为被其他人（如认知心理学家或人本主义心理学家）认为是由我们高度进化的意识或智力所致，这只是由于心理学家没能找到引发这一行为并使之维持的强化物。如果你觉得这是一个相当极端的观点，那么别忘了斯金纳的观点被称为“激进的行为主义”，这种观点一直饱受争议。

虽然经常遭到质疑，但斯金纳始终用实验支持自己的观点。他认为某些看似是人类独有的行为，也可以被像鸽子或老鼠这样的低等生物学会。例如，人们会认为迷信行为是人类所独有的，理由是迷信需要认知活动（即思考、辨别和推理）的参与。迷信是一种信仰，而我们通常认为信仰并不属于动物。但斯金纳提出，在本质上迷信行为与其他任何一种行为，很容易用操作性条件反射原理来解释。他设计了一个实验来证明这一观点。

理论假设

回想一下你的迷信行为，你是否曾轻敲木头以求好运，忌讳从梯子下走过，忌讳踩到裂缝，携带一枚幸运硬币或其他护身符，在棋盘游戏中用一种特定的方式掷骰子，因为算了命而改变自己的行为？可以肯定地说，尽管一些人不想承认，但在某些时候每个人都会因为迷信做过某种事情。斯金纳说，人们这样做的原因是他们相信或揣测在迷信行为和某些强化性的结果之间存在联系，虽然实际上两者并不相关。人们相信这种联系是因为该行为（如用某种方式摇骰子）被偶然地强化过（如掷出一个好点数）一次、两次或几次。斯金纳称它为非关联性强化（noncontigent），这种强化与特定行为间并不一一对应，但你却相信这种因果联系确实存在。“如果你认为这是一些人类特有的行为，”斯金纳说，“那么我将给你一只迷信的鸽子。”

方法

为了使读者理解斯金纳在这个实验中使用的方法，有必要对“斯金纳箱”做一个简短的介绍。斯金纳箱（斯金纳称它为条件反射箱）背后的原理其实相当简单。它由一只空的笼子或箱子组成，里边只有一个用来放食物的碟子或托盘。研究者可以用它来控制动物何时得到强化，如一些粒状食物。早期的斯金纳箱中还包括一根控制杆，如果按下控制杆，食物就会自动出现。如果把一只老鼠（斯金纳最早的研究中使用老鼠）放在这样的箱子里，老鼠会通过尝试错误和强化最终学会用按控制杆来获得食物。另一方面，如果需要的话，实验者可以通过控制食物分配器来强化某种特定的行为。后来，斯金纳和其他研究者发现鸽子也是条件反射实验中的理想被试，于是操作箱内设计了啄击盘子来取代控制压杆。

在本研究中，斯金纳使用了这样的操作箱，但是做了一个重要的改变。为了研究迷信行为，食物分发器被设定为每隔15秒落下食丸，不管动物当时在做什么。我们可以看到这便产生了非关联性强化。换言之，不管动物做了什么，每隔15秒它将得到一份奖赏。

研究中的被试是8只鸽子。连续几天对这些鸽子喂少于它们正常进食量的食物，以便让它们在测试时处于饥饿状态，由此增强其寻找食物的动机（这增加了强化的效果）。让每只鸽子每天在实验箱里待几分钟，对其行为不作任何限制。在这期间，每隔15秒强化自动出现。几天后，两个独立的观测者记录了鸽子在箱中的行为。

结果

斯金纳在报告中写道：

8只鸽子中有6只产生了非常明显的反应，两名观察者得到了完全一致的记录。一只鸽子形成了在箱子中逆时针转圈的条件反射，在两次强化之间转二到三圈；另一只

反复将头伸向箱子上方的一个角落；第三只显现出一种上举反应，好像把头放在了一根看不见的杆下面并反复抬起它。还有两只鸽子的头和身体呈现出一种钟摆似的动作，它们头部前伸，从右边快速地摆动到左边，接着再慢慢地返回，它们的身子也顺势移动，动作幅度过大时还会踉跄几步。还有一只鸽子形成了不完整的啄击或轻触的条件反应，动作直冲地面但并不触及地面（P.168）。

这些行为在建立条件反射前未曾观测到。实际上，新行为与食物的出现毫无关系。然而，这些鸽子表现得好像某种特定的行为会产生食物，也就是说，它们变得迷信了。

接下来，斯金纳想知道如果两次强化之间的时间间隔被拉长，又会发生什么。他用了一只摇头的鸽子，然后把两次投放食丸间的间隔慢慢增加到1分钟。这时，鸽子表现得更加精力充沛，直到最后在两次强化间的一分钟内，这只鸽子像在表演一种舞蹈（好像一种“鸽子用餐舞”）。

最后是消除鸽子的这种新行为。这意味着在测试箱中的强化不再继续。这时，迷信行为逐渐消退，直到完全消失。然而，值得注意的是，这只“跳舞”的鸽子在完全消退前的反应次数超过了10000次！

讨论

在本研究中，斯金纳得到了6只迷信的鸽子。然而，他对这一成果的解释却极为仔细和谨慎：“这一实验可以说证明了一种迷信。鸽子表现得好像其行为和食物间存在因果联系，虽然这种联系实际上并不存在。”（P.171）

下一步是要将这些结果应用于人类。你可能想到了人类行为中的类似情况，斯金纳也同样想到了。他写道：“打保龄球的人扔出球去，球沿着球道滚动，但他还在弯曲着胳膊和转动肩膀，似乎可以用此来控制球。”（P.171）从理性角度来讲，这些动作并不能真正对从球道滚到中途的保龄球起作用。但是由于过去建立的条件反射，你相信这些动作是有作用的。实际上，球投出去之后，无论它会滚到哪儿，跟你的行为已经没有关系。正如斯金纳所说：“掷球者的行为对球并没有影响，但球的运动对掷球者却有反作用。”（P.171）换言之，有时候球可能碰巧沿着掷球者身体移动的方向运动，球的运动及其全击中或丝毫未击中的结果偶然强化了掷球者弯曲或转动的行为，并使这种迷信得以保持。这与斯金纳的鸽子有区别吗？确实没有。

迷信难以消退的原因可以从那只在行为消除前跳了10000多次舞的鸽子中找到。当某种行为在某种特定情境中只是偶然地被强化一次（称为部分强化），它就变得非常难以消除。这是因为人们的期望值很高，期望迷信行为可以产生强化的结果。设想一下，如果这种联系每次都出现，然后突然消失，那么行为就会很快停止。然而，在现实生活中，意

外强化的事件通常是非常偶然地发生的，因此迷信行为常常会持续一生。

批评和后续研究

斯金纳的行为主义理论和研究总会引起广泛且激烈的争论。其他有关人类行为的主要流派主张，以及严格的行为主义观点不能解释人类的很多基本心理过程。卡尔·罗杰斯（Carl Rogers）是人本主义心理学的奠基人，因与斯金纳的论战而出名，他将人们对行为主义的批评总结如下：

> 在内心世界中，人本主义心理学可以研究那些对于行为主义者而言毫无意义的内容，如意向、目标、价值、抉择、自我知觉、对他人的知觉以及个人建构自我世界的结构基础……个体的整个现象世界及其意义的联结体。这个世界没有一个方面是向严格的行为主义者敞开的，但以上这些因素对人类行为有重要意义，这似乎是肯定的。（罗杰斯，1964，P.119）

行为主义者会反驳说，所有这些人类特性都可以用行为主义的方法进行分析。问题的关键是，对构成这些特性的行为和结果作出适当的解释（参见斯金纳，1974，对这些问题的全面讨论）。

然而，单就迷信这个问题而言，出现的争议却很少，并且由此形成的学习过程理论也获得了相当广泛的认可。由布鲁纳和列维斯基（Bruner & Revuski，1961）设计的一个实验证明了人类是多么容易形成迷信行为。4名高中学生，每人坐在4个电报机按键前。他们被告知，每当按到正确的键时，都会出现铃声并闪烁红灯，同时他们将得到5美分（相当于现在的50美分）。正确的反应是3号键。然而就像在斯金纳的实验中一样，只有在10秒的延迟间隔后，3号键产生的强化物才会出现。在这10秒内，学生会尝试各种组合的按键。接着，在延迟后的某一时刻，他们会得到强化物。对所有的学生来说，结果都是相同的：一段时间后，他们都各自形成了一套按键顺序，并在每次强化之间一遍一遍重复（如1，2，4，3，1，2，4，3）。其实，按下3号键才是唯一被强化的行为，而其他的按键序列完全是迷信。他们不仅迷信地行动，而且所有学生都相信按其他键是“建立”强化键所必需的，只是他们完全没有意识到自己的迷信行为。

近期应用

斯金纳是心理学界最有影响的人物之一，他在很多方面依旧对科研领域产生着深远而重要的影响。他于1948年发表的关于迷信行为的文章每年都会被无数的研究所引用。例如，有一项研究比较了两种强化类型在迷信形成中的作用（Aeschleman，Rosen，& Williams，2003）。行

为的后果是得到一些你想要的东西（例如金钱、食物或赞扬），属于正强化；负强化是通过消除掉你不喜欢的事物来进行奖赏（例如可以不做家庭作业或避免疼痛）。该研究发现，更高水平的迷信行为（对非偶然事件的有意控制）是在负强化下的条件下形成的，而非正强化下产生的。用作者的话来说，就是“这些结果……显示，相对于正强化，负强化对迷信行为的形成和维持提供了一个更有利的条件”（P.37）。换句话说，你更可能采取迷信策略去防止坏结果的发生，而非创造好的结果。

另一篇发人深省的文章也引用了斯金纳1948年的研究（Sagvolden et al.，1998），这篇文章考察了强化对注意缺陷性障碍/多动症（ADHD）的作用。研究者请患注意缺陷性障碍/多动症的男孩和没有患这种病的男孩共同参与一项活动，在游戏中，他们可以获得硬币或小玩具作为奖励。虽然强化固定为每30秒出现一次（非关联性强化），但所有男孩都产生了自认为与奖励有关的迷信行为。在下一阶段的研究中，强化不再出现。研究者希望以此减少和停止各种条件反射行为（即消退）。这种结果在没有注意缺陷性障碍/多动症的男孩身上确实出现了，但对于注意缺陷性障碍/多动症男孩来说，经过一个短暂的停顿后，他们却变得更加活跃，并且开始以一种突发冲动的方式作出反应，好像强化又再次出现了。作者指出，这种过度活跃和冲动意味着注意缺陷性障碍/多动症男孩与对照组男孩相比，应对强化延迟的能力明显不足。这些研究结果对我们深入认识和有效治疗注意缺陷性障碍/多动症是非常重要的。

结论

迷信无处不在。你可能就有一些，而且你也一定认识其他有迷信行为的人。有些迷信是文化的一部分，它们产生着广泛的社会影响。你可能会发现大部分高层建筑没有第13层。其实，事实并不是那样。很显然第13层一定存在，但没有任何一层楼叫“13”层。这并不是因为建筑师和建造者是一群极端迷信的人，而是由于13是个不吉利的数字，要租出或卖掉13层的房子十分困难。另一个例子是美国人对2美元纸币十分迷信，致使财政部发行的2美元纸币比其他任何面值的都少（不到1%）！

迷信从心理学角度看是不健康的吗？绝大部分心理学家相信，尽管从定义上讲，迷信行为并不会导致你想要的结果，但它们还是有积极的功能的。当一个人身处困境时，迷信行为经常能产生一种力量感和控制感。有趣的是，从事危险职业的人比其他人更加迷信。有时候，由迷信行为带来的力量感和控制感能降低人们的焦虑，增强其信心和勇气，进而提高成绩。

Aeschleman, S., Rosen, C., & Williams, M.(2003).The effect of non-contingent negative and positive reinforcement operations on the acquisition of superstitions behaviors.Behavioral Processes, 61, 37-45

Bruner, A., & Revuski, S.(1961).Collateral behavior in humans.Journal of the Experimental Analysis of Behavior, 4, 349-350.

Rogers, C.R.(1964).Toward a science of the person.In F.W.Wann (Ed.), Behaviorism and phenomenology: Contrasting bases for modern psychology.Chicago: Phoenix Books.

Sagvolden, T., Aase, H., Zeiner, P., & Berger, D.(1998).Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder.Behavioral Brain Research, 94(1), 61-71.

Schneider, S., & Morris, E.(1987).The history of the term radical behaviorism: From Watson to Skinner.Behavior Analyst , 10(1), 27-39.

Skinner, B.F.(1974).About behaviorism.New York: Knopf.

研究12 观察到攻击行为……做出攻击行为

Bandura, A., Ross, D., & Ross, S.A.(1961).Transmission of aggression through imitation of aggressive models.Journal of Abnormal and Social Psychology, 63, 575-582.

攻击性行为表现为各种各样的形式，是目前美国和世界面临的最大社会问题。同时，它也是心理学史上一个最重要的研究课题之一。多年以来，在行为科学家当中，一直处于这项研究最前列的是社会心理学家，他们研究的重心是人与人之间的相互作用。社会心理学家的目标之一是给攻击行为下一个准确的定义。这看起来似乎很简单，但实际上给出一个这样的定义相当困难。比如，你觉得下面的哪些行为是攻击行为：一场拳击赛？猫咬死一只老鼠？士兵向敌人开枪？在地下室安放捕鼠夹？斗牛？这些模棱两可的例子不计其数。因此，如果请教10位不同的社会心理学家，可能就会得到10种不同的攻击行为定义。

很多研究者已不再纠缠于给攻击行为下定义，而是转向考察人类攻击性的来源这一更重要的方面。他们经常提出的问题是：为什么人们会做出攻击性行为？纵观心理学的历史，人们已经提出很多理论假设来解释攻击性行为产生的原因。有些理论认为，人先天性地表现出攻击行为，因为在某些情况下攻击性行为可以说是生物进化的一种生存机制。比如，长期形成的暴力冲动不断累积，最后终于爆发。另一些理论强调情境因素，例如，多次受挫或受到特别的挑衅是引起攻击性行为的主要决定因素。第三种同时也是本研究提出的观点，认为攻击行为是习得的。

心理学史上有一个非常著名、非常有影响力的实验，它阐述了儿童是怎样习得攻击行为的。这项研究是阿尔伯特·班杜拉（Albert Bandura）和他的助手多萝西娅·罗斯（Dorothea Ross）以及希拉·罗斯（Sheila Ross）于1961年在斯坦福大学完成的。班杜拉被视为“社会学习理论”心理学派的奠基人之一。社会学习理论家认为，人与人之间的交互作用是人格发展的主要因素，比如，在你的成长过程中，父母和老师等重要人物强化了你的某一行为而忽视或惩罚你的其他行为。班杜拉认为除直接的鼓励和惩罚之外，行为的塑造还有一种重要的方式，即可以通过简单地观察、模仿（或以别人为榜样）其他人的行为而形成。

从本项研究的题目可以看出，班杜拉、多萝西娅·罗斯以及希拉·罗斯已经证明了榜样对攻击行为的影响。这项广为人知的研究在心理学界被称为“波比娃娃”研究，稍后我们将对它展开讨论。以往的参考文献中提到，早期研究结果证明儿童很容易模仿作为榜样的成人的行为。在这项新研究中，班杜拉想探讨在没有榜样出现的情境中，孩子是否会将这种

模仿学习泛化进去。

理论假设

研究者计划让儿童分别观察两名成人，一名表现出攻击行为，另一名不表现出攻击行为，随后在没有榜样出现的新情境中对儿童进行测试，以了解儿童在多大程度上模仿他们观察到的成人攻击行为。依照这种实验操作，班杜拉和他的助手们作出了四种预测。

1.观察到攻击行为的被试不论榜样是否在场，都会模仿成人做出类似的攻击行为。而且这种行为明显不同于观察到非攻击行为或根本没有榜样的被试。

2.对于观察到非攻击行为的儿童，他们的攻击性不仅比观察攻击行为的儿童更低，而且也明显低于无榜样的控制组儿童。换言之，非攻击性榜样能起到抑制攻击行为的作用。

3.因为儿童倾向于认同父母或与自己同性别的其他成人，被试“模仿同性榜样的行为远远超过异性榜样的行为”（P.575）。

4.由于在社会上，攻击行为主要是一种极典型的男性行为，男孩比女孩更倾向于模仿攻击行为，尤其是在给被试呈现男性榜样时差异更明显（P.575）。

方法

这篇文章条理清晰地概括了实验中所用的方法。本文作了适当的省略和简化，具体步骤如下。

被试

研究者得到了斯坦福大学附属幼儿园的管理人员和教师的支持和帮助，从而获得研究中所需要的被试。参加这项研究的被试由36名男孩和36名女孩组成，他们的年龄在3至6岁之间，平均年龄为4岁零4个月。

实验条件

24名儿童被安排在控制组，他们将不接触任何榜样；其余的48名被试先被分成两组：一组接触攻击性榜样，另一组接触非攻击性榜样，随后再按男女分组。最后，各组分出一半被试接触同性榜样，另一半接触异性榜样。这样最终得到8个实验组和1个控制组。你可能会问这样一个问题：如果某些儿童原先就比其他人更有攻击性怎么办？班杜拉通过事先获得每个被试的攻击性评定等级，从而克服了这种潜在的问题。一名实验者和一名教师（都是对这些儿童非常了解的）对这些儿童的身体攻

击、语言攻击和对物体的攻击行为进行评定。这些评定结果使实验者可以依据平均攻击水平对各组被试进行匹配。

实验程序

每个儿童分别单独接触不同的实验程序。首先，实验者把一名儿童带入一间活动室。在路上，实验者假装意外地遇到成人榜样，并邀请他过来“参加一个游戏”。儿童坐在房间的一角，面前的桌子上有很多有趣的东西。有土豆印章（这是在1961年，生长在高科技时代的你们可能没有见过，土豆印章就是把土豆从中间切开，在上面进行雕刻，就像橡皮图章，按过印泥后可以复制出很多几何图形）和一些贴纸，这些贴纸颜色非常鲜艳，还印有动物和花卉，儿童可以把它们贴在一块贴板上。随后，成人榜样被带到房间另一角落的一张桌子前，桌上有一套儿童拼图玩具，一根木槌和一个5英尺高的充气波比娃娃（一种大的不倒翁充气小丑，当用拳猛击或用脚踢时能再竖起来）。实验者解释说这些玩具是给成人榜样玩的，然后便离开房间。

无论是在攻击情境，还是在非攻击情境中，榜样一开始都先装配拼图玩具。但一分钟后，攻击性情景下的榜样开始用暴力击打波比娃娃。对于在攻击条件下的所有被试，榜样攻击行为的顺序是完全一致的：

> 榜样把波比娃娃放在地上，然后坐在它身上，并且反复击打它的鼻子。随后榜样把波比娃娃竖起来，捡起木槌击打它的头部，然后猛地把它抛向空中，并在房间里踢来踢去。这一攻击行为按以上顺序重复三次，中间伴有攻击性的语言，比如“打他的鼻子……打倒他……把他扔起来……踢他……”和两句没有攻击性的话：“他还没受够。”“他真是个顽强的家伙。”（P.576）

这样的情况持续将近10分钟，然后实验者回到房间里，向榜样告别后，把孩子带到另一间活动室。

在无攻击行为的情境中，榜样只是认真地玩10分钟拼图玩具，完全不理波比娃娃。班杜拉和他的同事们努力确保除要研究的因素——攻击性榜样对非攻击性榜样以及榜样性别——以外的所有实验因素对每名被试都是一样的。

愤怒或挫折感的激发

10分钟的游戏以后，在各种情境中的所有被试都被带到另一个房间，那里有非常吸引人的玩具，如救火车模型、喷气式飞机模型、包括玩具木橱和玩具车在内的一整套玩具娃娃，等等。研究者相信，为了测试被试的攻击性反应，使儿童变得愤怒或有挫折感会令这些行为有更多发生的可能。为了实现这种目的，实验者先让被试玩这些有吸引力的玩具，不久以后，告诉他们这些玩具是为其他儿童准备的。但是也告诉被试，他们可以到另一间房间里去玩别的玩具。

检测对攻击行为的模仿

在最后的实验房间内，有各种攻击性和非攻击性的玩具。攻击性玩具包括波比娃娃（当然会有）、一个木槌、两支掷镖枪和一个上面画有人脸的系绳球。非攻击性玩具包括一套茶具、各种蜡笔和纸、一个球、两个玩具娃娃、小汽车和小卡车，以及塑料动物。允许每个被试在这个房间里玩20分钟，在这期间，评定者在单向玻璃后依据多条指标对每个被试行为的攻击性进行评定。

攻击行为的评定指标

总共评定被试行为中的8种不同反应。为清楚起见，在此，我们只概述4种最鲜明的反应：首先，研究者记录所有对榜样的攻击行为的模仿，包括坐在波比娃娃身上，击打它的鼻子，用木槌击打它，用脚踢它，把它抛向空中；第二，评定被试对攻击性语言的模仿，记录他重复“打他”“打倒他”和“砰”等的次数；第三，记录被试用木槌进行的其他攻击行为（也就是用木槌击打娃娃以外的其他东西）；第四，用列表的方式列出成人榜样未做出而被试自发做出的身体或语言的攻击行为。

结果

观察者把研究发现概括为表3—3。如果你仔细查看此表，就会发现这些结果支持了班杜拉和罗斯等人在实验前提出的四种假设中的三种。

那些看到榜样做出暴力行为的被试倾向于模仿这种暴力行为，男性被试每人平均有38.2次、女性被试平均有12.7次模仿了榜样的身体攻击行为。此外，男性被试平均17次、女性被试平均有15.7次模仿了榜样的言语攻击行为。这些特定的身体和言语攻击行为，在无攻击行为榜样组和控制组几乎没有发现。

表3—3 儿童在不同处理条件下攻击反应的平均数

攻击类型	榜样类型				
	攻击性男性	无攻击性男性	攻击性女性	无攻击性女性	控制组
模仿身体攻击					
男孩	25.8	1.5	12.4	0.2	1.2
女孩	7.2	0.0	5.5	2.5	2.0
模仿言语攻击					
男孩	12.7	0.0	4.3	1.1	1.7
女孩	2.0	0.0	13.7	0.3	0.7
用木槌攻击					
男孩	28.8	6.7	15.5	18.7	13.5
女孩	18.7	0.5	17.2	0.5	13.1
自发攻击行为					
男孩	36.7	22.3	16.2	26.1	24.6
女孩	8.4	1.4	21.3	7.2	6.1

（引自P.579）

回想一下，班杜拉和他的助手曾预测无攻击行为的榜样能对儿童的攻击行为起到抑制作用。支持这种假设的结果应该显示被试在无攻击条件下攻击行为的平均数明显低于没有榜样的控制组。在表3—3中，如果你对比一下非攻击性榜样组的数字和控制组的数字，你会发现结果是混杂的。比如，在木槌攻击行为中，观察无攻击行为男性榜样的男孩和女孩表现出的攻击行为明显低于控制组；而观察无攻击行为女性榜样的男孩表现出的攻击行为却远远高于控制组。作者承认，这种矛盾性结果不能说明无攻击榜样能对攻击行为产生抑制作用。

实验假设中提到的性别差异从表3—3数据中得到了有力支持。很显然，男孩受男性攻击性行为榜样的影响明显超过同样条件下的女性榜样。观察男性榜样的攻击行为后，男孩每人平均表现出104次攻击行为，而观察女性榜样后，平均只有48.4次。另一方面，女孩的行为虽然不太一致，但观察女性榜样的攻击行为后，平均出现57.7次攻击行为，而观察男性榜样后，只有36.3次表现出这种行为。作者指出，在同性别模仿条件下，女孩更多地模仿言语攻击，而男孩更多地模仿身体攻击。

最后，几乎在所有条件下，男孩比女孩都更明显地表现出身体攻击的倾向。如果把表3—3中的所有攻击行为的数据相加，男孩总共表现出270次暴力行为，女孩则只有128次。

讨论

班杜拉和罗斯称，他们已经证明在不给榜样或观察者任何强化物时，特定行为（这里指暴力行为）是如何通过观察和模仿而习得的。由此得出以下结论：成人的行为向儿童传递了这样一个信息，即这种形式的暴力行为是允许的，这样便削弱了儿童对攻击行为的抑制。他们指

出，当儿童以后遇到挫折时，他们可能更容易表现出攻击行为。

研究者同时探讨了为什么攻击性的男性榜样对男孩的影响明显大于女性榜样对女孩的影响。他们解释道，在美国文化（以及世界大部分国家的文化）中，攻击行为被看成是典型的男性行为，而不是女性行为。换言之，它是一种男性化的行为。所以，攻击性的男性榜样带有更大的社会认可度，因此可能对观察者的影响更大。

后续研究

在实验进行的时候，研究者们可能还没想到它将来会产生多大的影响。20世纪60年代初期，在美国文化中，电视已经成为一种重要力量，消费者们开始逐渐认识到电视暴力对儿童的影响。对于这个问题，一直存在着激烈的争论。在过去的30年中，美国国会至少举行了3次有关电视暴力的听证会，班杜拉和其他心理学家的工作也被涵盖其中。

两年后，班杜拉和他的那两位助手进行了一项后续研究，目的是考察电视或其他非人类的攻击榜样对被试的影响力。他们使用了类似于“波比娃娃研究”的实验方法。班杜拉、罗斯和罗斯设计了一个实验，比较了真人榜样和电影或卡通片中的榜样做出相同攻击行为时对被试的影响。结果证明，真人榜样影响力最大，电影榜样位居第二，卡通片中的榜样名列第三。但不管怎么说，三种形式的攻击性榜样对儿童的影响都比非攻击性榜样或无榜样明显得多（Bandura，Ross & Ross，1963）。

令人振奋的是，班杜拉在后来的研究中发现，在特定的条件下，榜样的暴力影响可以被改变。回想一下他早期的研究，实验者并未对榜样和被试的攻击行为给予奖赏。但你是否想过，当儿童观察到榜样因暴力行为而得到强化或惩罚时又会怎么样？班杜拉（1965）检验了这种观点，并发现当儿童看到暴力行为受奖励时，他们会更多地模仿暴力行为；当榜样的暴力行为受惩罚时，他们会明显减少对攻击行为的模仿。

对班杜拉这些研究的批评指出，对充气娃娃的攻击行为不同于对人的攻击，并且孩子们知道这两者的不同。在班杜拉和同事研究的基础上，其他研究者检验了榜样的暴力行为对真正攻击行为的影响。在一项研究中，研究者使用了班杜拉的波比娃娃的实验方法（Hanratty，O’Neil，& Sulzer，1972），先让孩子们观察成人榜样的暴力行为，然后使他们产生很大的挫折感。此时，孩子们常常会对真人（一个穿得像小丑的真人）产生暴力行为，不论这个人是否是引起挫折感的原因。

近期应用

本章所讨论的班杜拉的研究对心理学至少有两方面的主要贡献：首先，它在很大程度上说明了儿童的新行为是怎样通过简单地模仿成人而习得的，甚至成人可以不真正出现，社会学习理论家认为，构成一个人人格的许多行为（即便不是绝大多数）都是通过模仿形成的；第二，在过去45年间，这项研究为数百项关于现实生活中的暴力或媒体暴力对儿童影响的研究奠定了基础（若要了解班杜拉及其对心理学的贡献，可参见Pajares，2004）。

在不到10年的时间，美国国会又举行了多次听证会，以探讨包括电视、电影、视频游戏、电脑游戏和互联网对人的巨大潜在影响。广播和多媒体开发者越来越多地感受到公众和法律带来的压力，他们正在努力减少媒体暴力，或者设立“家长忠告”等级体系，对暴力内容给予特别提醒。

也许更加令人担忧的是，越来越多的证据表明，媒体暴力对儿童的影响效果会持续到成年期（Heusmann et al.，2003）。一项研究发现“儿童期暴露于媒体暴力可以预见其青年期的攻击性行为，对于男性和女性均如此。对暴力电视节目中人物的认同和对电视暴力现实性的知觉，也能预见其后的攻击性。即使当社会经济地位、智力和教养方式等因素的影响被控制的情况下，这种相关依然存在”（Heusmann et al.2003，P.201）。

结论

由于儿童越来越容易接触到各种各样迅速发展的媒体渠道，社会对媒体暴力效应的关注也在不断地增加。因为阻止儿童接触到媒体暴力是不可能的，所以，研究的焦点不断集中到如何防止儿童将媒体暴力转化为现实攻击性的策略上。在美国的校园枪击案发生后，这方面的研究急剧增加，并且在可预期的将来会继续成为研究的前沿。最近，加利福尼亚议会通过了一项法令：禁止向未得到父母允许的18岁以下儿童出售“极端暴力”视频游戏，违法者处以1000美元罚金。你可能会问：“什么是极端暴力呢？”根据该法令，极端暴力被定义为对人造成严重伤害，特别是通过邪恶的、残暴的、残酷的手段实施的（参见Going after video game violence，2006）。如果你感到这一定义过于主观也不奇怪。视频游戏行业正在以违反宪法规定进行起诉，希望撤销该项法令。可以确信，班杜拉的研究将成为这场战役的一部分。

Bandura, A.(1965).Influence of models’ reinforcement contingencies on the acquisition of imitative responses.Journal of Personality and Social Psychology，1, 589-595.

Bandura, A., Ross, D., & Ross, S.(1963).Imitation of film mediated aggressive models.Journal of Abnormal and Social Psychology, 66,

3-11.

Going after video game violence.(2006).State Legislatures 32(1), 9.

Hanratty, M., O’Neil, E., & Sulzer, J.(1972).The effect of frustration on the imitation of aggression.Journal of Personality and Social Psychology, 21, 30-34.http：//www.ship.edu/~cgboeree/ bandura.Html (1998).

Huesmann, L.R., Moise, J., Podolski, G.P., & Eron, L.D (2003)Longitudinal relations between childhood Exposure to media violence and adult aggression and violence:1977–1992.Developmental psychology, 39(2), 201–221.

Pajares, F.(2004).Albert Bandura: Biographical sketch.Retrieved March 10, 2007, from Emory University, Division of Education Studies Web site: http// des.emory.edu/mfp/baandurabio.html.

第四章 智力、认知和记忆

在心理学的众多分支中，认知心理学是与本章主题关系最密切的一个分支。认知心理学研究人的心理过程。人的智力、复杂的思维和推理能力，以及储存和提取过去经历符号表征的能力，所有这一切构成了人与动物在本质上的区别。当然，这些能力极大地影响着人们的行为。然而，研究心理过程远比研究可以观察到的行为要困难得多，所以在研究中必须采用大量富有创造力的、灵活的方法。

本章收录的几项研究改变了心理学家对人类内部心理行为方式的看法。第一个研究讨论了著名的“皮格马利翁研究”，该研究表明，他人比如教师的期望不仅能影响学生在学校的表现，而且还能影响他们的IQ分数；第二篇文章讨论的研究改变了我们对人类另一个基本属性的感知，即我们的智能。20世纪80年代早期，霍华德·加德纳（Howard Gardner）提出人类不是拥有一种普遍的智能，而是至少拥有七种独特的智能。他的观点就是现在大家熟知的多元智能理论（MI理论）；第三项研究是认知心理学领域中的另一项早期研究，它检验了动物和人怎样对周围环境形成心理表象，这种心理表象被称为“认知地图”；第四项是一项新近研究，它揭示了人们的记忆并不像想象中的那样准确，它引起了我们对目击者法庭证词可靠性的思考。

研究13 期望导致结果

Rosenthal, R., & Jacobson, L.(1966).Teachers’ expectancies: Determinates of pupils’ IQ gains.Psychological Reports，19，115-118.

我们对于自我实现预言都很熟悉。这一概念的表现方式之一就是，如果我们预期某一事物将以某种方式发生，我们的期望就会倾向于让它变为现实。自我实现预言是否在生活中以预言的方式真实存在，这还有待于人们的科学研究，但是心理学研究已经证明，在某些领域，这的确是事实。

1911年，对“聪明的汉斯”的著名研究（Pfungst，1911）首次引起了心理学家对自我实现预言这一问题的关注。“聪明的汉斯”是冯·奥斯登（von Osten）先生的马，它因为能够认字、拼写、解决数学问题而远近闻名。在解决数学问题时，它用前掌击地的次数代表问题的答案。对此，自然会有很多怀疑者，但专家组对它的能力进行测试后发现：没有冯·奥斯登先生的提示，马也能真实地表现出这些能力。但是，马怎么会有相当于人的智力程度呢（也许20世纪60年代的著名电视喜剧里的Mr.Ed除外）？20世纪初期的心理学家奥斯卡·波菲格斯特（Oskar Pfungst）进行了一系列仔细的实验研究后发现，汉斯实际上是从提问者无意识的表现中获得细微线索的。例如，人们问完一个问题以后，一般都会低头俯视其前掌等待它的答案。当它前掌敲击的次数接近正确答案时，提问者就会微微地抬起眼睛或头，期待着马能完成它的解答。汉斯已对提问者的这些细微动作形成了条件反射，它会利用这种线索，并得出正确回答。

接着，你可能会问“这匹‘诡计多端’的马怎么会与心理学的研究有关呢？”对“聪明的汉斯”的研究提出了一种可能性，即观察者经常带有某种特殊的期望或偏爱，这就使他们在研究中无意识地给被试发出某种隐蔽的信号。这些信号就可能使被试按照与观察者的偏爱相一致的方式作出反应，从而证实了观察者的预期。所有这一切归根结底就是，实验者可能总是期望，在经过他或她的科学处理后，一名或一组被试与其他人相比就会产生某种特殊的行为。有时，这些行为只不过是实验者自己有倾向性的期望所导致的结果，除此之外，别无其他。如果事情果真如此，那么实验便是无效的。这种对心理学实验效度的影响被称为“实验者期望效应”（experimenter expectancy effect）。

研究这一方法论问题的先驱罗伯特·罗森塔尔（Robert Rosenthal）用心理学实验证明了实验者期望效应的存在。他利用正在进行《学习与条件反射》课程学习的心理系学生进行了一项实验研究（Rosenthal & Fode，1963），在此期间，这些学生并不知道自己已经成为了被试。罗

森塔尔告诉其中的一部分学生他们将要使用的是经过特殊方式喂养的有较高智力水平的“聪明鼠”，它们能快速学会走迷津；而告诉另一些学生他们将要使用的老鼠在学习走迷津时比较迟钝。然后，学生们开始训练老鼠完成各种任务，其中包括走迷津。与得到愚笨鼠的被试相比，得到聪明鼠的被试报告他们的老鼠所需的学习时间明显较短。事实上，分配给两组学生的那些老鼠都是标准的实验用鼠，而且这些老鼠都是随机分配的。这些学生并不是在说谎或是故意歪曲实验结果。他们在训练动物时，对动物施加的影响显然是无意识的。

一系列的实验结果说明，实验者期望效应对科学研究的威胁显然是存在的。训练有素的研究者会使用更严谨的实验程序来避免类似的期望效应（例如使用“双盲”实验，在这种实验中，与被试接触的实验人员并不知道研究假设）。

此外，罗森塔尔还谈到在实验室以外，类似的偏见和期望也可能会出现，如在学校的课堂中。因为那些公立学校的教师一般不会有机会了解期望的潜在危害，也不了解它会对学生的潜能产生多大的影响。毕竟在传统教育中，教师在一年级时就可以拿到学生的IQ分数。这些信息会不会使教师对不同学生产生带有倾向性的期望，并因此使他们无意识地区别对待“聪明”学生和所谓“不太聪明”的学生（以IQ分数为依据）？如果事实果真如此，那么这公平吗？这些问题构成了罗森塔尔和雅各布森的研究基础。

理论假设

当这种期望效应出现在实验室以外、人与人之间的自然交往中时，罗森塔尔称之为“皮格马利翁效应”（pygmalion effect）。在希腊神话中，雕塑家皮格马利翁爱上了他创作的女雕像。我们更熟悉的是现代剧作家萧伯纳的作品《皮格马利翁》（其音乐剧版本名为《窈窕淑女》），由于亨利·赫吉斯的教育、鼓励和期望，伊利沙·多利特（Eliza Doolittle）才得以才华出众。罗森塔尔怀疑，在小学教师得到学生的某种信息（例如IQ分数）时，他们或多或少地会对学生的潜能产生某种期望。这种期望会使他们无意识地对那些可能会成功的学生的行为表现给予一些鼓励和鞭策，使这些学生产生自我实现的预期，变得更加出色。当然，这似乎是以牺牲那些教师对其期望不高的学生为代价的。为了检验这些理论假设，罗森塔尔和他的助手雅各布森与一所小学（橡树学校）取得合作，这所小学位于某个大城镇的中低阶层生活区。

方法

开学初，在橡树学校工作人员的配合下，研究者对1到6年级的所有

学生进行了IQ测验（一般能力测验，简称TOGA）。选择这个测验的原因在于，它属于非文字测验，学生的分数不依赖于先前接受的阅读、写作和算术技能等方面的训练，而且橡树学校的教师对此项测验并不熟悉。研究者告诉教师，学生们所接受的是“哈佛应变能力测验”。在此情况下，这种隐瞒很有必要，其目的是让教师对学生产生一些期望，而这正是该实验成功的必要因素。研究者还进一步对教师解释道，该测验的成绩可以对一名学生未来在学术上是否会有成就作出预测。换言之，他们是要让教师相信在测验中获得高分的学生，其学习能力在未来的这个学年中将有所提高。实际上，这个测验并不具备这种预测能力。

在橡树学校，总共有六个年级，每个年级有三个班，共18名班主任老师（16名女教师，2名男教师）。每位班主任都得到了一份名单，上面记录着本班在哈佛测验上得分最高的前20%的学生，以便老师们了解在本学年里哪些学生有发展潜力。但是，下面才是本研究的关键：即教师所得名单中的前十名学生是被完全随机地分配到这种实验条件下的。这些学生和其他学生（控制组）的唯一区别就是，教师以为他们（实验组学生）会有不同寻常的智力得分发展。

接近学年结束时，研究者对所有学生再次进行了相同的IQ测验（即TOGA），并计算出每个学生IQ的变化程度。通过对实验组和控制组的IQ变化差异的检验就可以看出，在现实情境中是否也存在期望效应。

结果

图4—1显示了实验组和控制组学生的IQ提高情况。综合全校的情况来看，那些被教师以为智力发展会有显著进步的学生，其IQ平均提高幅度显著高于控制组的学生（分别为12.2分和8.2分）。然而，对图4—1进行仔细观察，我们就会发现，这种差异主要是由一二年级组中的巨大差异引起的。稍后，我们将讨论导致这种差异的可能原因。罗森塔尔和雅各布森提供了另一种更清楚有效地展示一二年级数据的方法，见图4—2。图4—2向我们展示了每组学生中IQ成绩分别提高了10分、20分、30分的人数比例。

在这项早期的研究中，研究者得到了两个主要的研究发现：一是已在正式实验室情境中被证明了的期望效应，也会在非正式的现实生活情境中起作用；二是这些作用在低年级中表现得更明显，而在高年级中几乎不存在。这又都意味着什么呢？

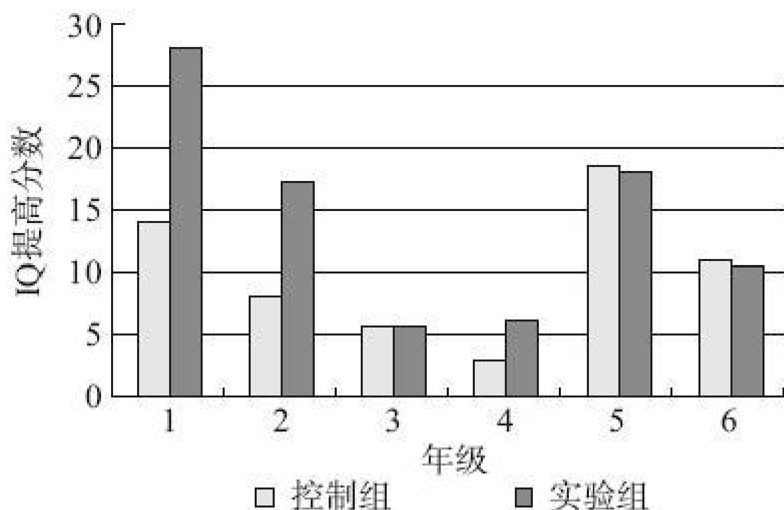


图4—1 1~6年级学生IQ分数增长图

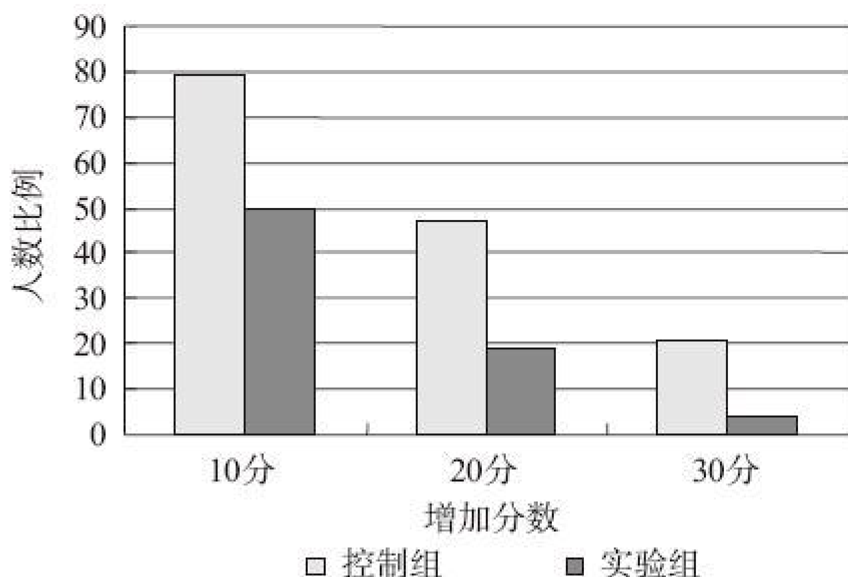


图4—2 一二年级学生IQ分数增加的人数的百分数

讨论

正如罗森塔尔在他以前的研究中所猜测的，教师对学生行为的期望转化成了学生的自我实现预言。“当教师期望某个孩子表现出较大程度的智力提高时，这名学生就真的出现了较大程度的智力提高”（Rosenthal & Jacobson，1968，P.85）。请注意，图4—2中报告的数据是每个年级

三名教师所教的三个班级的平均成绩。很难想象除了教师的期望以外，还有什么其他理由可以解释学生在IQ分数提高量上的差异。

此外，对高年级学生而言，自我实现预言的作用似乎并不明显。罗森塔尔认为对这一现象进行解释是很重要的。在本文和一些后续论文中，罗森塔尔和雅各布森对此提出了如下几种可能的解释。

1.低年级儿童的可塑性一般较高年级儿童更强。如果事实的确如此，那么研究中低年级儿童变化更大也许仅仅是由于他们比高年级儿童更易变化。与此相关的另一种可能性就是，即使低年级儿童并不具有很强的可塑性，但教师也会认为他们是有这种特性的。可能仅仅是这种想法，就足以让教师对学生采取不同的处理方式，从而产生不同的结果。

2.小学低年级学生还未能在老师的心目中形成牢固的印象。换言之，如果教师没有对学生的能力形成某种认识，那么该研究提出的期望就会产生更重要的影响。

3.低年级的学生可能更容易受到教师将期望传递给他们时的微妙过程的影响。

> 根据这种解释，如果教师相信某些学生能获得智力上的提高，那么她对待各年级的这类学生的方式可能是相同的。但也许只有低年级儿童的成绩会受到教师对待他们的特殊方式的影响，这些方式包括她对他们所说的一些话，说话时特殊的语气，她的眼神、姿势和与学生的身体接触等（Rosenthal & Jacobson，1968，P.83）。

4.低年级教师向学生传递期望的方式与高年级教师不同。如果确实存在这种情况，那这些差异可能会怎样？罗森塔尔和雅各布森没有对此作出推测。

研究发现的意义和后续研究

罗森塔尔和雅各布森在橡树学校的研究结果的真正意义在于，它表明教师的期望效应对学生的在校表现会产生长期的潜在影响。它直接引出了在当今心理学和教育学中非常有争议的一个话题：即IQ测验的公平性问题。稍后，让我们再回到这个问题上进行一番讨论，但首先我们要来探讨几项有趣的后续研究。这些研究考察了教师是如何在不经意间向他们认为有较大潜能的学生传递他们的期望的。

在查肯、西格尔和德来岗（Chaiken，Sigler & Derlega，1974）的研究中，研究者对课堂教学情境中的师生互动情况进行了录像，研究者告诉教师其班里的某些学生极为聪明（这些所谓的“聪明”学生其实是从全体学生中随机抽取的）。仔细观看录像，我们就会发现在很多细微之处，教师都表现出对“聪明”学生的偏爱。他们给予这些学生更多微笑、更多的眼神交流，对这些学生的课堂回答给予更多的赞同。研究者还指出，这些在教师心目中存在较高期望值的学生更喜欢学校生活，更乐意接受教师对其错误的建设性批评，并更努力地对此进行改进和提高。该项研究及其他一些研究的结果表明，虽然教师对学生的期望不是学生在校表现的唯一决定因素，但它的影响绝不仅限于IQ成绩。

假设你是一名小学教师，你的班级中有20名学生。在开学的第一天，你收到了印有每个学生IQ成绩的名册。你发现这些学生中有5人IQ成绩超过145分，可以归入天才的行列。你认为你会像对待其他学生那样对待他们吗？你认为你对他们的期望会与你对其他学生完全一样吗？你对他们的期望与你对另外5名IQ成绩低于正常值的学生的期望会有不同吗？如果你的回答是你对他们的处理和期望将完全相同，罗森塔尔可能会打赌你是错的。事实上，不可能完全相同！关键是如果你的期望变成了他们的自我实现预言，那么这对有些学生来说是不公平的。现在我们来考虑一下另外一点，也是更重要的一点：假设名册上的IQ成绩是错的，错误分数产生的期望有利于某些学生而不利于其他学生，这显然是不公平的，可能也是不道德的。这是当今激起人们对IQ测验进行争论的主要因素之一。

很多年来，对IQ测验始终存在着这样一种指责，即用于评价学生智力的标准化IQ测验中包含着种族和文化偏见。不少人认为，这些测验的设计者是中上阶层的白种男性，因此，测验中包含了一些其他种族的人们未曾接触过的思想和信息。在美国，少数民族儿童在这些测验上的得分往往要低于白人儿童。如果我们由此得出这些孩子与白人孩子相比智商较低的结论，那将是非常可笑的。所以，他们之间的分数差异只可能源于测验本身。但依照传统做法，自幼儿园开始直至12年级，教师们总能得到所有学生的IQ成绩。如果你仔细想想这一与罗森塔尔和雅各布森的研究相关的现象，就可能会发现潜在的危险情境正在出现。除了在学校中学生们可能会被教师按照IQ成绩分类（有提高班、补习班等）外，教师会以这些可能带有偏见的信息为依据，下意识地对学生产生某种期望，从而使学生产生不当的自我实现预言。支持这种观念的论据很有说服力，这使美国的大多数州都延缓使用IQ测验成绩；只有发明出的新测试方法（或将旧的方法升级）被证实是有效的且不存在偏见后，测验分数才能得以使用。本节所讨论的这项研究就是支持以上观点的经典论据。

近期应用

由于罗森塔尔和雅各布森的研究工作，教师期望效应对学生表现的影响已经成为我们在理解教育过程时的一个必不可少的部分。此外，罗森塔尔的“人际期望”理论已在除教育以外的其他领域产生了巨大影响。2002年，罗森塔尔本人利用元分析的方法对有关期望效应的文献进行了回顾（关于元分析的解释，见第9章史密斯和格拉斯的研究）。他向我们说明了“心理学研究者、学校教师、法官、企业主管、医疗服务提供者的期望如何能够在不经意间影响被试、学生、陪审团、雇员和患者的反应”（Rosenthal，2002，P.839）。

有一篇观点颇为犀利的论文引用了罗森塔尔的期望效应研究，研究者考察了教师把学生送到心理学家那儿接受心理评估和心理咨询的问题

（Andrews，Wisniewski & Mulick，1997）。研究者发现，教师让班级中非裔美国学生去进行发展性障碍评估的比率明显高于班里的白人学生。另外，教师把在教室里发生的问题以及在运动场上的行为问题更多地归咎于男生，较少归咎于女生，其间相距悬殊。研究者认为，不同组别的学生所表现的差异更多地是由教师期望的不同所致，而不是个体之间真实存在的差异。

值得注意的是，心理学界和教育学界的许多研究人员正努力研究新的方法，用于概念化和测量儿童的智力水平。一些研究带头人聚焦于现有的关于人类大脑的工作原理，并提出不同的测试方法。这些方法与以往简单地将一般智力分数命名为IQ的陈旧的、局限性的方法大不相同（参考Benson，2003），其中最具代表性的是罗伯特·斯滕伯格（Robert Sternberg）的三元智力测验（1993）。该测验可用于测量智能的三个显著方面：分析智能、实践智能和创造智能。另一个智力研究领域的代表人物霍华德·加德纳（Howard Gardner），早在20世纪80年代早期就提出了多元智能理论，该理论至今仍在智力研究和测量领域发挥着重要作用。正如你在下一节将会读到的加德纳的理论，他指出，我们不仅仅有一种或三种智能，而是有八种不同的智能（也可能有九种或更多），而且每个人的不同智能的发展水平也是不一样的（Gardner，2006）。

Andrews, T., Wisniewski, J., & Mulick, J.(1997).Variables influencing teachers’ decisions to refer children for school psychological assessment services.Psychology in Schools, 34(3), 239-224.

Benson, E.(2003).Intelligent intelligence testing: Psychologists are broadening the concept of intelligence and how to test it[Electronic version].Monitor on Psychology, 34(2), 48.

Chaiken, A., Sigler, E., & Derlega, V.(1974).Nonverbal mediators of teacher expectancy effects.Journal of Personality and Social Psychology, 30，144-149.

Gardner, H.(2006).Multiple intelligences: New horizons.Jackson, TN: Perseus Books Group.

Pfungst, O.(1911).Clever Hans (the horse of Mr.von Osten): A contribution to experimental, animal, and human psychology.New York: Holt, Rinehart and Winston.

Rosenthal, R.(2002).Covert communication in classroom, clinics, courtrooms, and cubicles.American Psychologist, 57, 839-849.

Rosenthal, R., & Fode, K.(1963).The effect of experimenter bias on the performance of the albino rat.Behavioral Science, 8，183-189.

Rosenthal, R., & Jacobson, L.(1968).Pygmalion in the classroom: Teacher expectations and pupils' intellectual development.New York: Holt, Rinehart and Winston.

Sternberg, R.J.(1993).Sternberg Triarchic Abilities Test.Unpublished test, Yale University.

研究14 你在哪方面更聪明

Gardner, H.(1983).Frames of mind: The theory of multiple intelligences.New York: Basic Books

本研究的标题是作者有意玩的一个文字游戏。这个问题通常的问法是“你有多聪明”，这意味着你的智力是多少（即从“量”的角度来衡量）。本研究的标题是问“你在哪方面更聪明”（即从“质”的角度来衡量），这就和“量”毫无关系了，而实质上是关于某种特定的智力类型。这意味着，人并不是简单地表现为聪明或不聪明的，我们每个人都拥有一个独特的、各种组合形式的智力能力。

许多人都至少做过一次智力测验（即使你想不起来了），有些人可能做过更多。在过去一百多年以来开发的智力测验中，绝大部分都是以一个分数表示结果，这个分数被称为“智力商数”（IQ）。如果智力测验的结果是一个分数，那么智力的概念必然就是人类的一种单一的、一般的心理能力。纵观整个20世纪，大多数时候，智力正是这样被解释的。人们经常用字母“g”（general）来表示一般智力或能力。当时在各种环境中，包括学校、工厂、军队等，IQ分数（即g）都作为一项标准被广泛地用于对人的评价、分类以及描述上。

然而，在20世纪七八十年代，研究者们开始质疑用这种单一的、一般的智力理论方法来测试人类智力的合理性。大多数智力测验本身就对特定的经济阶层和种族带有偏见，而且儿童受教育的机会也经常取决于这种带有偏见、甚至可能是不合理的IQ分数（见第13项研究中罗伯特·罗森塔尔针对这种偏见带来的危害所举的例子）。

随着对早期智力概念的批评逐渐增多，且影响不断扩大，IQ智力测验使用得越来越少。而此时，一种崭新而又截然不同的智力观点科学且通俗地解释了智力的作用。与一元化的智力理论完全不同的是，这种新兴理论把智力概念扩展为许多不同类型的智能，而每种智力都有各自完整而又独立的特征。哈佛大学的霍华德·加德纳（Howard Gardner）于1983年在其著作《智能的结构》（Frames of Mind）中向世界介绍了多元智能理论，这正是本节研究的基础。

理论假设

加德纳多元智能理论（The theory of multiple intelligences，MI）不仅仅基于对人们表现出来的形形色色的相关智力技能的简单观察，他

的理论来源于对大脑结构的研究。实际上，在着手研究智能之前，加德纳的主要工作是研究生物学和大脑的功能。他进一步拓展了前人研究的成果，证明人类大脑不仅仅在能力上有所区别，而且在功能上有着明确的分工。换句话说，经过进化后的不同脑区负责执行与思维和认知有关的特定任务。脑功能定位理论主要采用脑损伤法，即通过观察大脑某个区域受损害的患者，其在那方面的能力衰退或消失，以此表明该脑区所具有的功能。加德纳同样采用了该研究方法。例如，大多数人的语言功能位于大脑左半球的一个区域；视觉功能处于大脑后侧枕叶皮层的中心；而位于视觉皮层底部的一个特殊结构则具有辨认和识别人脸的功能（更多关于大脑功能定位的理论见本书第1章迈克尔·加扎尼加的割裂脑研究）。

随着大脑功能定位理论的进一步发展，加德纳认为人类大脑的不同区域负责智力的不同方面，或者更准确地说，就是不同的智能。为了增强其多元智能理论的科学性，加德纳对很多不同类型的人进行了测试，而且他还制定出将一组技能界定为一种特定智能类别的标准。加德纳是这样描述其研究资料的：

为了能系统阐述多元智能理论，我以大量不同类型且迄今毫无关联的人作为考察对象。这些人包括：神童、天才、大脑损伤的病患、白痴天才（idiot-savants，大脑迟钝的一种罕见类型或是在一个或两个领域有天赋的自闭症患者）、普通儿童、普通成人、不同行业的专家以及有着不同文化背景的人。（P.9）

方法

综合这些研究数据，加德纳制定了八项界定智能的标准或“标志”。如果某种智力或某一组能力被认为是独立存在的一种智能，那么它必定会符合以下绝大多数标准。

1.大脑损伤证明智能具有潜在独立性。加德纳认为一种特殊能力如果伴随着大脑损伤而受到破坏，或者当其受损时其他能力仍能保持完整，那么就有充分的理由相信这种能力是一种独立的智能。

2.白痴天才、神童以及天才的存在都与独立智能有关。你会发现一部分人在某个特殊能力上拥有超常的智能；一些大脑迟钝或自闭的人被证明是“天才”；而有些智力平平的人其实是神童，他们拥有的某些能力远远超出他们的实际年龄和经历范围。加德纳相信这些个体的特殊能力为证明独立智能的存在提供了有力的支持。

3.智能应该与一个清晰的信息加工（思维）过程相联系。这就意味着智能是一种特殊的思维能力。因为如果一种能力有资格被界定为智能，它就必须包含一系列特殊的智能加工过程，加德纳称之为“核心操作系统”（core operations）。它们存在于大脑的特定区域并且只会对某种特定信息作出反应。表4—1列举了加德纳所界定

的各种智能的“核心操作”。

表4—1 加德纳八种智能的“核心操作系统”与典型例证

智能	核心操作	典型例证
语言	句法学（词的组合）、语音学（声音的发声）、语义学（词的意思）、语用学（词的使用）	莎士比亚、J. K. 罗琳、苏斯博士、伍迪·艾伦
音乐	音调（声音的频率）、节奏，音色（声音的质量）	莫扎特、格温·史蒂芬妮、安德里亚·波切利、保罗·西蒙
逻辑-数学	数字、数量、范畴、因果关系	阿尔伯特·爱因斯坦、卡尔·萨根、玛丽·居里、B. F. 斯金纳
空间	精确的视觉、表象旋转和转换	毕加索、弗兰克·劳埃德·赖特、莱昂纳多·达芬奇、文森特·梵高
肢体-运动	对自己身体的控制，对目标的掌握	查理·卓别林、勒布朗·詹姆斯、塞莉娜·威廉姆斯、维纳斯·威廉姆斯
人际	对他人的感觉、感情、目标和动机的意识	莫汉达斯·甘地、亚伯拉罕·马斯洛、奥普拉·温弗瑞
自我内省	对自身的感觉、感情、目标和动机的意识	柏拉图、赫尔曼·罗夏、海伦·凯勒、比尔·盖茨
自然观察	对自然环境中的事物的认知和分类；对自然世界的敏感性	查尔斯·达尔文、简·古道尔、雷切尔·卡逊
存在*	从事于先验性的关注的能力，如人类生存的基础、人生意义和死亡意义	埃利·维瑟尔、小马丁·路德·金、小卡尔·罗杰、伊丽莎白·库伯勒·罗斯

*候选。

4.智能应该有一个独特的发育历史和发展到高级专业水平的潜力。加德纳相信每种智能都应该经历一条发展的道路：始于简单，逐步前进，并越过一个又一个象征能力更高等级的里程碑。

5.智能的逐步进化发展有据可依。人类智能作为一种适应机制已经经历了数百万年的进化，并使人类得以生存。根据跨文化研究的结果以及对动物身上相似能力的观察，加德纳坚信，如果一种特殊能力被界定为智能，那么其涉及的技能应该显示出进化发展的迹象（例如，本书第15个研究中关于托尔曼老鼠“认知地图”的研究）。

6.智能可以用心理学实验进行研究。加德纳认为任何被界定为智能的能力都必须能够采用严密的实验来证明。例如，图形空间旋转的反应速度和准确度的实验可以用来测量一个人空间智能的高低。图4—3就是这项实验的一个例子。你能多快反应出来呢？

7.智能可以采用现存的标准化测验来测量。这里，加德纳肯定了IQ及其他智力测验的可能价值。然而他所看到的价值并不是这些测验能产生一个关于智力的分数，而是一些测验所包含的子维度，可以用来测量不同的智能。

8.智能的各个类型都可能用一个符号系统代表。最后，加德纳提出任何类型的人类智能都应该整合为一个符号系统。最明显的当

然是人类的语言和数学智能。其他的如代表音乐智能的乐谱、代表空间智能的图画都是符号系统典型的例子。

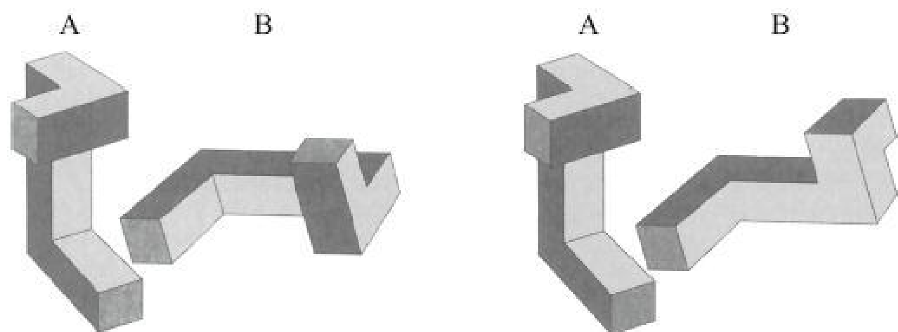


图4—3 用图形空间旋转实验测试空间智能的例子

注：两幅图中的图形是相同的还是不同的？

接下来，我们将看到加德纳在其1983年出版的著作中对多元智能所作的总结。他所涉及的每种智能都按照其制定的八项标准进行了分析。不符合多数标准的能力就会被排除。经过这一淘汰过程，加德纳最初提出了七种截然不同的智能，后来增加到八种，最近又提出第九种智能。

结果

加德纳在1983年的著作中，对其提出的七种智能进行了详尽的讨论。在这里，我们对每一种智能进行了简要的描述，而且我们还引用了一些加德纳文章中的论述，旨在使读者对这些能力有更清晰的理解。此外，表4—1总结了每种智能的核心操作系统，并列举了一些在这一智能类型上能力很强的著名人士。尽管加德纳并不认同任何一项单一的测验可以测出多种智能，但还是开发出了一些这样的测验。你可以去做做，只要在网上搜索“多元智能测试”就可以找到，但要记住网上的很多材料的效度是有待商榷的。

语言智能

如果你有突出的语言智能（linguistic intelligence），在使用文字上，你会比一般人更熟练、更恰当、更富有创造性。你可以用语言说服别人支持你的立场；你可以记住并能回忆起详细或复杂的信息；你比绝大多数人都更能擅长解释和传授概念、想法等；而且，你也在充分享受着运用语言来谈论语言本身所带来的乐趣。加德纳认为有天赋的诗人就是语言智能突出的代表：

在诗人对某行或某段诗歌字斟句酌的过程中，我们可以发现语言智能的某些方面在发挥着作用。诗人必须对词义的细微差别极度敏感，并且还要尽可能多地使其保留下来并广为流传……具有对词语使用规范的敏感性，遵守语法规则或是合理选择时机去违背规则的能力。如果语言智能达到更高的层次——对发音、节奏、变音以及韵律都很敏感，那么这样的诗人创作出来的诗歌，甚至外国人诵读起来也会觉得悦耳动听。（P.77~78）

音乐智能

你可能已经猜出了音乐智能（musical intelligence）的某些构成要素：包括对声音，特别是声调、音色以及节奏的天赋能力。加德纳声称音乐智能是所有智能里最早出现的，音乐神童为“音乐天才”提供了例证。加德纳用作曲家的例子来解说音乐智能：

一个作曲家最明显的特征就是在其大脑中“不断有音乐出现”，也就是说，他经常能在意识表层附近听到曲调、节奏以及更大一些的音乐形态。（P.101）

逻辑-数学智能

逻辑–数学智能（logical-mathematical intelligence）可以使你思考、分析、推断抽象的事物、概念或观点之间各种复杂的关系。数学家、科学家和哲学家往往拥有高水平的逻辑-数学智能，但是，这也可能出现在那些爱好运动分析、电脑编程以及开发算法的人们身上：

这类人的典型特征是偏爱解决抽象性问题……数学家必须绝对地严谨并长期持有怀疑态度：任何事实只有经过一系列严密步骤证明后才能被认可，而这些步骤也必须是在那些普遍真理的指导下推导出来的……很明显，他们快乐来源就是成功地解决长期未能被解决的问题。（P.138–141）

空间智能

如果你在脑中形成表象、想象以及操纵表象的能力优于常人的话，就说明你拥有很高的空间智能（spatial intelligence）。这些能力都是与生俱来的。它们在那些从事依靠想象力的工作或以其为业余爱好的人们身上很容易找到，例如艺术家、雕刻家、室内装饰设计人员、工程师和建筑师等。为了使人们对空间智能了解得更具体，加德纳进行了如下解释：

辨认相同结构的能力；转换或识别结构变形的能力；在脑中想象并转换图像的能力；根据空间信息绘制图表的能力；以及其他类似的能力。（P.176）

图4—3关于图形空间旋转的测试对于拥有高水平空间智能的人来说一点都不难。

肢体运动智能

肢体运动智能（bodily kinesthetic intelligence）也可以被称为“身体智能”。如果你在这方面水平很高，那么你会对自己的身体和身体的运动非常敏感，同时，你也可以灵活自如地使用和控制你的身体去实现各种目标。舞蹈家、运动员、外科医生、制陶工人和很多演员都拥有很高的身体运动智能。加德纳进一步解释说：

> 这种智能的特点是为了实现对充满表现力和目的性目标的追求，能够十分娴熟地使用高度分化的身体部位……另一个特点是拥有灵活操纵物体的能力，包括那些靠手指或手进行的精巧运动和那些靠整个身体进行的大型运动。（P.206~207）

接下来，加德纳提出的两种智能，尽管相互独立，但却属于同一类型。加德纳称之为“个人智能”（personal intelligences）。在这两种智能中，一种关注人的内部世界，而另一种则关注外部世界。它们就是内省智能和人际智能。

内省智能

你对自己有多了解？加德纳指出，人们意识和了解自己以及自己的情感、动机和行动的根源存在不同程度差异的能力。加德纳对高水平的内省智能（intrapersonal intelligence）是这样描述的：

> 这种智能的核心能力是对自身感情生活的了解——一个人的情感领域。它能使人快速辨别出这些感情，最后对它们进行分类，将其转换为象征性的符号代码，并利用这些代码来了解和指导自身的行为。（P.239）

人际智能

该智能与内省智能相对立，它关注的是“你对其他人有多了解”。人际智能（interpersonal intelligence）所包括的能力与内省智能差不多，但它们都是指向外部的，即关注的是对其他人的情感、动机、欲望和行为的了解：

> 这种智能的核心能力是对他人的察觉和区分，尤其是意识和区别他人的情绪、性格、动机、意图的能力。人际智能作为一种高级形式的智能，它使一个成熟的成年人能看出别人隐藏于内心的意图和欲望，并会据此采取相应的行动。（P.239）

以上这七种智能构成了加德纳最初的多元智能概念。但加德纳在《智能的结构》中明确指出：这些智能只是一个初步的罗列。经过进一步的调查与研究，很可能会纳入其他智能，原有的七种智能也可能由于某种具有说服力的观点会去掉一个或多个。许多年过去了，这七种智能仍保留了它们原来的地位，而且加德纳又增加了第八种（也许还有第九种）智能。

后续研究与批评

加德纳的多元智能理论（MI理论）迅速得到教育家、家长以及社会大众的推崇，因为它证明了一个大家普遍持有的信念：人的聪明可以表现在不同方面。这就解释了为什么那些在考试或某些学科表现欠佳的学生（也包括成人）却在其他方面表现得十分聪明。

MI理论被用到对学习障碍儿童的关注和研究上，在很大程度上促使教育界将“学习障碍”重新命名为“学习差异”。实际上，MI理论已经在教育领域发挥了极大的作用。加德纳在其著作《智能的结构》出版后所进行的研究聚焦于将其观点运用到改进儿童和成人的教育上。

在多元智能理论提出10年后，加德纳又对其理论进行了修订，并认为有其他方面的能力成为也可能被归为智能。于是，加德纳的同事向他建议了很多不同领域的能力，如“精神智能”、“性智能”和“数字智能”（Gardner，2003）。虽然加德纳承认将一组能力归为一种智能有不同的主观解释，但他始终认为这些及其他提议都未充分地满足那八项标准，因此，不算作是智能。然而，加德纳在研究中发现有一种能力非常符合其智能判断的标准。有一次，加德纳的一位同事让他描述历史上最有影响力的生物学家所具备的智能，正当他准备去描述时，却意识到在其定义的七种智能中没有一种是适合这些人的。这促使加德纳增加了第八种智能，即“自然观察智能（naturalist intelligence）”。加德纳解释说：

> 自然观察智能是指辨识和归类植物、矿物和动物的能力，包括岩石、草地以及各种各样的植物群和动物群。达尔文当属自然学家中最著名的典范，原因就在于他把生物的本质看得如此透彻。（Checkly，1997）

目前为止，介绍的这八种智能构成了加德纳的MI理论，但加德纳并没有停止他的理论研究。他认为多元智力概念是动态的：始终面向新的、有明确定义的能力。加德纳认为有一种能力非常符合他的智能标准，即存在智能（existential intelligence）。不过存在智能在十分接近MI理论标准的阈限，所以本文将其作为一种新的智能罗列在表4—1中。加德纳对存在智能的解释如下：

> 作为一种候选智能，存在智能立足于人类，倾向于思考关于存在这一最基本问题。我们为什么而活？我们为什么死去？我们从哪里来？将来会发生什么？什么是爱？为什么要发起战争？有时我觉得这些是超越感知的问题。他们关注的问题太大或太小，以至于都已经超出了我们的五官所能感知的能力。（加德纳，2006，P.20）

自1983年《智能的结构》一书出版后，加德纳又出版和发表了为数众多的著作和文章来完善其理论，并尝试将这些理论运用到相关的教育情境中。事实上，与其他学习或思维环境相比，MI理论已经更多地被运用在教育情境中，尤其是在幼儿园到高中阶段中的应用。例如，在《智能的结构》出版仅一年后，印第安纳波利斯的某重点学校就完全按照MI理论重新设计了课程。而今天，在美国及其他国家，几乎每所学校都在不同程度上运用着该理论。

尽管MI理论是关于人类智能研究中最负盛名的理论，并且已经在各

种研究和教育领域受到普遍支持，但它也不是无可挑剔的。在任何学科内，那些挑战传统观念且富有影响力的新生理论都会成为被激烈争论的焦点，MI理论也不例外。常见的一种反对观点指出，加德纳提出的八种智能并不能真正区分人类的智能，而仅仅是对人类不同的“思维方式”进行了描述，它们包含在本文开篇讨论的一元（g）智力理论中（Morgan，1996）。另一种批评认为多元智能理论存在很多内在矛盾，以至于过于含糊，难以证明其准确性（Klein，1998）。此外，克莱因（Klein）还认为，正是由于这种模棱两可，MI理论被当成一种“便利工具”来解释几乎所有的认知活动，使其根本不需要证明也不会受到反驳。另外，有些研究者的争论集中在两方面：一方面，认为加德纳没有进行足够严谨的科学研究来证明这些智能的准确性；另一方面，则质疑MI理论运用到实际情境中的有效性。他们强调，如果将来经过研究发现MI理论不是一个正确的或有效的工具，那么大量的时间和大部分努力就都白费了。同时对学习的影响在现实中并不会出现（Collins，1998）。尽管存在这样或那样的批评，MI理论仍在人类智能研究领域有着巨大的影响力。

近期应用

每年，都有数以百计的基于霍华德·加德纳的多元智能理论的科学文章和著作发表，且每篇文章和著作都引用他于1983年出版的那本书。加德纳博士在该领域的研究仍对学习与智能的研究及思考产生着重大且深远的影响。为了使读者了解MI理论的不同应用，接下来将简单介绍两项近期发表的相关研究。

首先，介绍一项关于加德纳七种智能的跨文化研究。该研究主要比较了英格兰和伊朗学生对自己及其家庭成员各项智能的评价（Furnham et al.，2002）。最主要的有趣结果是：与英格兰学生相比，（1）伊朗学生一般对自身逻辑—数学智能的评价较低，而对自己的空间智能、音乐智能和内省智能的评价较高；（2）伊朗学生对自己父亲的数学智能和空间智能的评价较低，但是对父亲的内省智能和人际智能评价较高；（3）伊朗学生认为自己的母亲除了内省智能外，其他智能都较低；（4）伊朗学生评价他们的兄弟在各项智能（除了数学智能）上都较高。

另一项令人振奋的研究是桑德拉·贝姆（Sandra Bem）的双性化研究（详见本书第26项研究中对贝姆研究的介绍）。研究发现，人们对自身智能的评价与其性别认同有关（Rammstedt & Rammsayer，2002）。研究者让被试对其自身各项智能进行评价，并同时用贝姆性别角色量表测定他们的男性化、女性化以及双性化的程度。经过分析发现，被试不仅在逻辑-数学智能（男性化）和音乐智能（女性化）上表现出了性别差异，而且男性对自我男性化、女性化或双性化的认同显著地影响着他们对自身各种智能水平的评价。

结论

加德纳的MI理论自提出至今已有20多年了，而且仍然焕发着旺盛的生命力。该理论是否会持续发挥重要影响力，还是会被新的智能理论所代替，我们将拭目以待。无论它将来发展得如何，有一点是肯定的：MI理论永远地改变了世界对学习、教育和智力的看法。然而，加德纳告诫我们，MI理论是达到目的一种手段，但它自身不应该被视为一种目的：

> 教育的目的应该能反映出一个人自身的价值，而这种目的不可能简单或直接地通过一种科学理论来实现。当一种科学理论能反映出个体的教育价值并能阐明个体的教育目标时，多元智能（假定存在的话）可以对该理论给予非常有效的证明。特别是当个体的教育目标包括对学科的理解时，就可以调用这几个智能去帮助实现这一高层次的目标……我已经认识到，当一个人向世界提出一种观点时，那他将很难控制这种观点的发展和影响，就像任何人很难控制自己基因的产物——孩子。简单地说，我最广为人知的思想——MI理论，现在和将来都将有它自己的道路，而且很有可能比我希望的更好。（Gardner，2002）

Checkley, K.(1997).The first seven and the eighth.Educational Leadership, 55, 8-13.

Collins, J.(1998).Seven kinds of smart, Time，152, 94-96.

Furnham, A., Shahidi, S., & Baluch, B.(2002).Sex and cultural differences in perceptions of estimated multiple intelligence for self and family: A British-Iranian comparison.Journal of Cross-Cultural Psychology, 33, 270-285.

Gardener, H.(2003).Multiple intelligences after twenty years.Paper presented at the American Educational Research Association, Chicago, IL, April, 21, 2003.

Gardner, H.(2006).Multiple intelligences: New horizons.Jackson, TN: Perseus Books Group.

Klein, P.(1998).A response to Howard Gardner: Falsifiability, empirical evidence, and pedagogical usefulness in educational psychologies.Canadian Journal of Education, 23，103-112.

Morgan, H.(1996).An analysis of Gardner's theory of multiple intelligence.Roeper Review，18, 263-269.

Rammstedt, B., & Rammsayer, T.(2002).Gender differences in self-estimated intelligence and their relation to gender-role orientation.European Journal of Personality，16, 369-382.

研究15 心中的地图

Tolman, E.C.(1948).Cognitive maps in rats and men.Psychological Review, 55，189-208.

有不少研究被收录于本书的原因在于，这些研究的理论假设和研究结果反驳了当时流行的看法和传统观点。例如，鲍查德关于人格受遗传影响的发现（研究3）、华生关于小艾尔伯特的研究（研究10）、哈罗的婴儿依恋理论（研究17）以及其他的一些研究，都对心理学思想的现状提出了挑战，从而开辟了对人类行为新的甚至是革命性的新解释。加利福尼亚大学伯克利分校的心理学家爱德华·C·托尔曼（Edward C.Tolman）关于学习和认知的理论探讨和研究正是作出了这样的贡献。多年来，心理学被严格的刺激—反应学习理论所统治，这一理论无视不能直接观察的内部心理活动的存在。托尔曼用实验证明，不仅对人类可以不用直接观察就能研究其内部心理过程，还可以对老鼠复杂的内部认知活动进行研究。鉴于其研究的重大意义，托尔曼被公认为“认知—行为主义”（cognitive-behaviorism）这一学习理论流派的奠基人。

为了更好地体会托尔曼的研究设想，让我们来举个例子，假设你现在想去最近的邮局或音像店，你很可能已在头脑中有了这些位置的图像。现在思考一下你要走的那条路线，你很清楚必须要走哪些街道，在哪儿转弯，并最终到达目的地。你当前的位置与邮局或音像店的路线关系在你头脑中所形成的这幅画面被称为“心理表征”（mental representation）。托尔曼则将这些表征称为“认知地图”（cognitive map）。他认为不仅人类使用认知地图，而且其他动物（包括老鼠）也在其思维活动中使用认知地图。为什么会有人关心老鼠的思维方式呢？因为对于20世纪三四十代的学习理论研究者来说，他们主要采用的研究方法就是让老鼠学习走迷宫，而令研究者非常感兴趣的是这些老鼠是如何学习的。

理论假设

在20世纪上半叶，学习理论家是走在心理学最前沿的，除了要解释学习机制，他们还致力于证明心理学的科学地位。由于心理学从哲学中独立出来而成为一门科学也仅有几十年的时间，很多研究者认为，要想证明心理学的科学性，最好的方法就是去效仿诸如物理学和化学那些所谓的“硬”科学。这种想法致使学习理论家指出，心理学正如物理和化学一样，它的唯一合适的研究对象是可观察和可测量的事件。作用于机体的某种刺激是可测量的，机体对刺激的反应行为也是可测量的，但在刺

激和反应之间的机体内部发生了什么，则是无法观察或测量的，因此，它不能被研究，而且就此而论也就不重要了。根据这种观点，当老鼠学习走迷宫的速度越来越快、错误越来越少时，这一学习过程就是由一连串刺激以及由老鼠在最后走出迷宫时的食物奖赏所引起的一连串正确反应而组成的。关于行为的这种严格的刺激-反应联结思想构成了行为主义的精髓，而且在行为心理学历史的前50年中一直占据着统治地位。

然而，在20世纪三四十年代，出现了一支以托尔曼为首的反对行为主义观点的派别，他们主张在学习过程中不仅有刺激和反应，而且在机体内部还发生了比这更复杂的过程。事实上，托尔曼对当时流行的观点提出了两点主要的修正意见。一是，如果不对与刺激和反应同时发生的内部心理过程进行考察，就不可能充分理解学习的本质及其复杂性。正如托尔曼在他1948年发表的著名论文中所阐述的（这篇论文也正是我们讨论的主题）：

我们相信在老鼠学习走迷宫的过程中，它的头脑中已建立了类似环境地图的东西。我们同意其他学派（刺激—反应）的观点，即老鼠在走迷宫时接受了某种刺激，并最终由于这些刺激而导致了实际中所出现的各种反应。不过，我们觉得，发生在大脑中的活动远比刺激-反应心理学家所认为的更复杂、更丰富多彩，而且通常也更自主。（P.192）

托尔曼提出的第二点是，尽管内部认知过程无法直接观察，但我们可以通过分析可观察的行为客观、科学地将其推断出来。

方法和结果

托尔曼在他1948年的论文中介绍了很多研究，以支持他的观点。所有这些研究都是关于老鼠学习走迷宫的实验。下面将介绍的两项研究简单明了地验证了托尔曼的理论观点。

第一个实验称为“潜伏学习”（latent learning）实验。实验中，把老鼠分为三组：对C组（控制组）而言，研究者所使用的是学习走迷宫的标准程序——让老鼠练习走迷宫，在迷宫的出口处放着食物作为奖励，每日如此；对N组（无奖励组）的老鼠，其每天被放入迷宫的时间与C组相同，但不出现食物，而且在迷宫中的任何行为都不会受到奖励；而D组（延迟奖励组）的老鼠在前10天受到与N组同等的待遇，但自第11天起，研究者会在迷宫的出口处放置食物，而且以后每天如此。

以三组老鼠的平均错误数（进入盲巷的次数）为指标，实验结果如图4—4所示。在该图中很容易看出，N组和D组的老鼠在没有得到任何奖励时，它们的学习没有多少进步；而控制组的老鼠在2周的时间内就可以达到近乎准确无误的程度。但是，当D组的老鼠一旦发觉走迷宫能得到好处（食物）时，它们仅用3天就近乎无误地学会了走迷宫（从第11天到第13天）。对结果的唯一可能解释是，该组老鼠在学习走迷宫的

前10天中，它们所学到的东西比它们表现出来的要多得多。正如托尔曼所解释的："一旦……它们知道能得到食物，它们的表现证实了在先前没有奖励的练习阶段，它们已经习得了许多盲巷的位置。它们已经构建了一幅'地图'，而一旦产生走出迷宫的动机，它们就会立刻利用这幅'地图'。"（P.195）

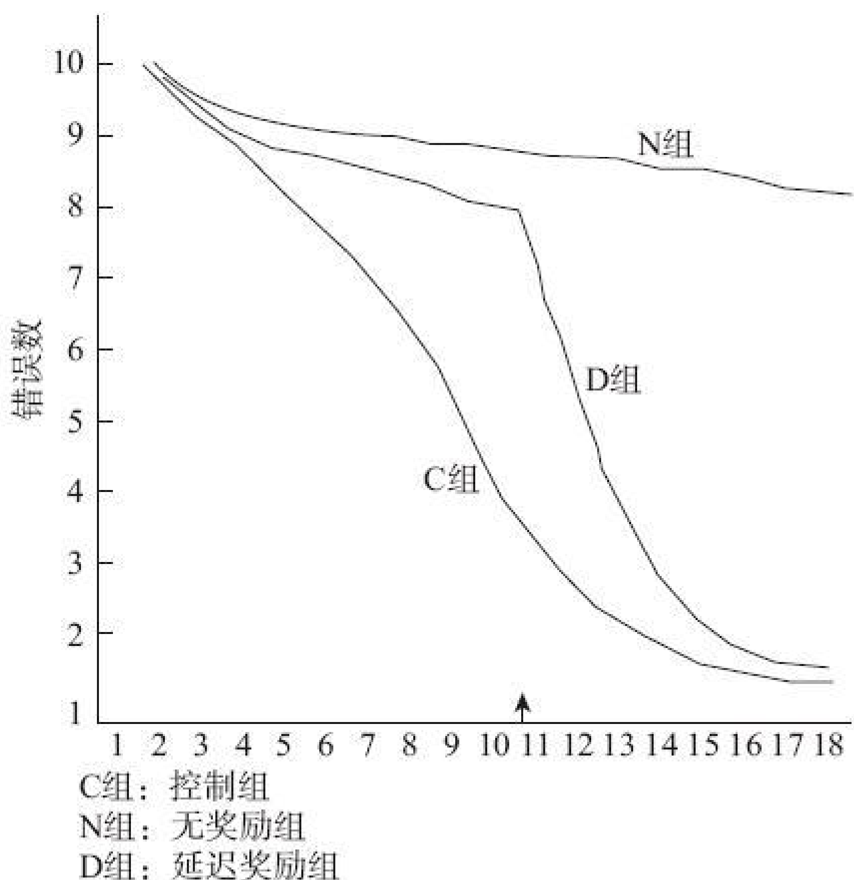


图4—4 在潜伏学习实验中老鼠学习走迷宫的错误率（引自P.195）

第二项研究是"空间定向"实验。刺激—反应理论认为，老鼠为了得到食物奖赏，只有通过在迷宫里四处奔跑，并经历过所有的刺激—反应联结（即S-R）后，才能知道食物的具体位置。这就好像你只有走出厨房，走过客厅，穿过走廊，经过洗手间，进入你的卧室，才能知道卧室的具体位置。实际上，你对卧室在家中的具体位置已形成了一种心理表征，而并不必像"走迷宫"那样来找到卧室的位置。托尔曼设计的空间定向技术就是要说明，进行走迷宫训练的老鼠实际上掌握的是食物出现的位置与它们的出发位置间的相对关系，而不管迷宫的结构是否发生了巨大变化、甚至被拆除。

首先，老鼠学习图4—5所示的简单迷宫。它们进入迷宫的入口，穿

过圆台面并进入引导通道，经过一条迂回曲折的路线，走到有食物奖励的出口。这是一个相对简单的迷宫，对老鼠而言，经过12次试验就能近乎准确无误地完成整个走迷宫任务。

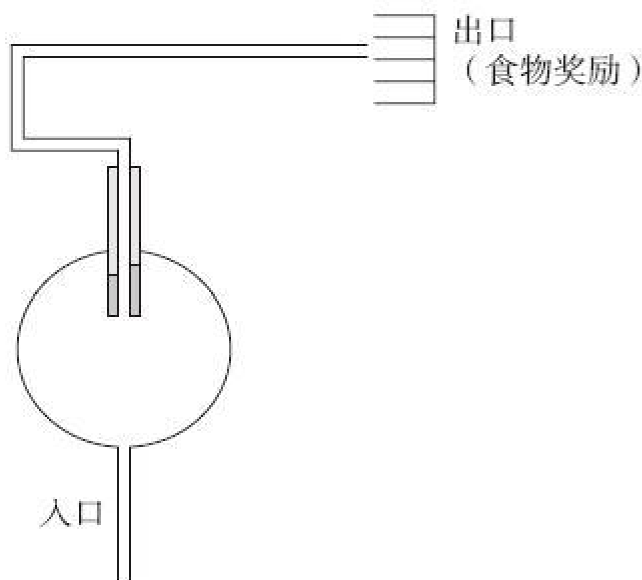


图4—5 空间定向实验：简单的迷宫

资料来源：引自P.202。

然后，把迷宫改成图4—6所示的放射状。现在，当受过训练的老鼠试图走它们过去的路线时，却发现道路被堵住了，它们只能回到圆台面上，并且必须在几条可能的备选路线中做出选择，以便到达先前放有食物的迷宫出口处。图4—7列出了选择各条路线中的老鼠数量。

正如你所看到的，与其他路线相比，老鼠更多地选择了出口距先前食物出现的位置仅4英寸的路线6。根据刺激—反应理论，老鼠最可能选择最接近先前迷宫中第一个转弯的方向的路线11，但事实并非如此。“看来老鼠掌握的不只是使它能按特定路线找到食物的序列地图，而是掌握了一幅含有食物的具体位置及其在房间内的具体方位的更广泛的综合性地图”（P.204）。在此，托尔曼把他得自老鼠的认知地图理论（即老鼠可以产生从A点到Z点路线的认知地图）推广到包括人类在内的其他有可能的生物有机体上。他证明了有机体所形成的认知地图不是一种从A点到B点到C点……再到Z点的序列地图，而是一幅更广泛、更综合化或概念化的地图，它使生物有机体在大脑中形成了一种认知“地貌”（lay of the lamd）。

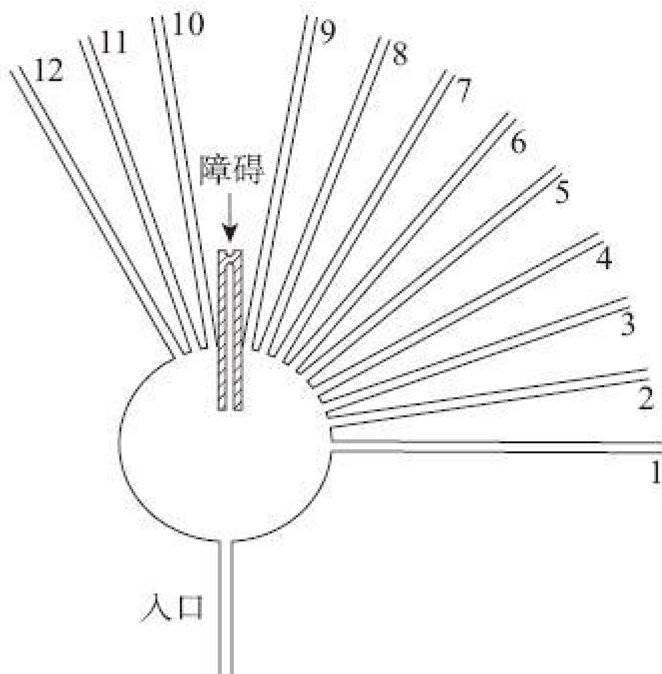


图4—6 空间定向实验：放射状的迷宫

资料来源：引自P.203。

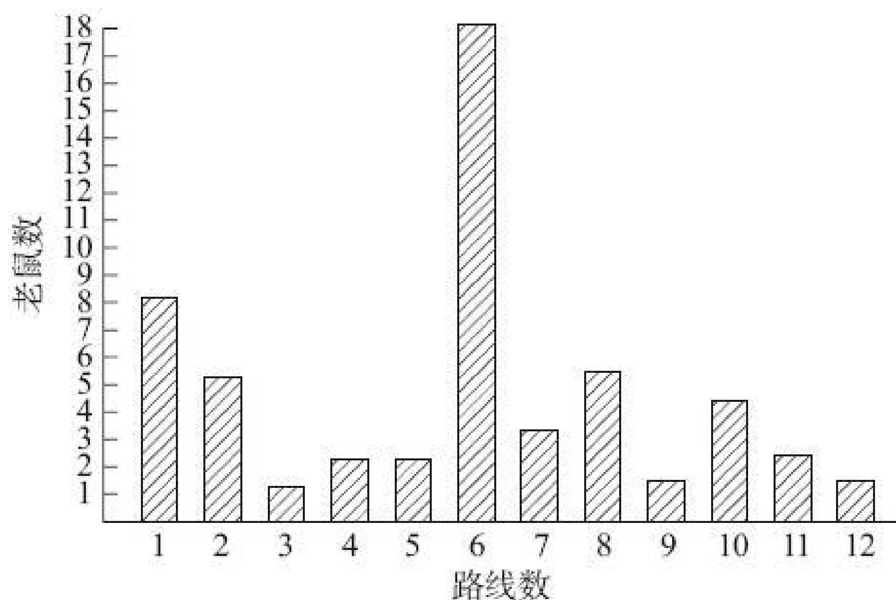


图4—7 空间定向实验：选择各条路线的老鼠数量

资料来源：引自P.204。

讨论

托尔曼在其1948年论文的结论性评述中，集中探讨了内涵单一的序列地图与更广泛综合的认知地图间的差异。当托尔曼把这一发现应用于人类时，他构建了这样的理论：因社会环境所形成的综合地图对人类而言是有益的，而内涵单一的序列地图则可能会使人陷入消极状态之中，如心理疾病或偏见和歧视等。他的推论建立在与上述研究结果有关的基础之上。他发现，当老鼠有过分强烈的动机（如过于饥饿）或者受到过多挫折（例如盲巷太多）时，它们倾向于形成内涵单一的序列地图，而不是如前所述的综合认知地图。托尔曼承认他不是临床心理学家或社会心理学家，但他仍把他的上述观点看作是对一些社会问题的一种合理解释。托尔曼论述道：

过分强烈的动机或极度的压抑状态使人类一而再、再而三地误入盲目仇视外来者的歧途。他们对外来者的仇视表现在方方面面，从对少数民族的歧视到世界大战的爆发，形式多种多样。

以上帝或心理学的名义，我们该做些什么呢？我唯一的答案就是重申理性的力量，也就是综合的认知地图……我们不该让自己或其他人过于情绪化、过于饥饿、过于衣衫褴褛，动机过于强烈，这些只能形成狭隘的序列地图。我们所有人……必须保持平和的心态，吸收充足的营养，以便形成真正的综合性的认知地图……简言之，当我们的孩子或我们自己面对人类世界这一上帝赐予的“大迷宫”时，我们必须使我们的孩子和我们自己（就像善良的实验者对待他的实验老鼠那样）处于能激发适度动机而没有多余挫败感的最理想的状态中（P.208）。

后续研究及近期应用

在托尔曼完成其早期研究之后的几十年里，大量的研究结果都支持了他的认知学习理论。现在，认知心理学已经成为当今行为科学中最具活力和影响力的研究领域之一，这也许是托尔曼的思想观点最著名的成果。作为心理学的一个分支，认知心理学主要研究无法观察的内部认知过程，而在几十年以前，“心理”这一概念还被排除在科学的研究对象之外。但如今人们对心理学的态度却有了一个近乎180度的大转变，即人们已普遍接受了这样的观点，即在决定一种行为反应的过程中，经由感觉、注意、思维、期望、记忆和分析对刺激进行的心理加工至少与刺激本身的作用等同，甚至可能更重要。

托尔曼的认知地图理论还影响了心理学的另一个分支——环境心理学。该领域关注的是人类行为与行为发生的环境之间的关系。环境心理学中的一项重要研究内容是研究人们怎样体验和看待生活中的各种环境，如城市、邻里、校园或办公楼等。人们对地点形成概念化的认识被称为“环境认知”，而对这些地点形成的精确心理表征正是托尔曼所谓

的“认知地图”。环境心理学家借鉴托尔曼的基本概念，不仅在了解人们对环境的理解方面，而且在规划环境或者使环境与我们的认知地图加工过程达到最佳匹配方面，都取得了颇具影响力的研究成果。

将托尔曼的思想应用于人类的环境心理学家之一便是林奇（Lynch，1960）。林奇提出了五种人们用来形成认知地图的环境要素。道路被认为是承载交通的主干道，无论是乘车、步行、骑自行车或乘船均如此。边界是我们在认知地图中划分一个区域与另一个区域的分界线，但其功能与道路不同，它可以是峡谷、一堵墙或某个湖的岸边等。交汇点则是一些引人注目的地方，是道路和边界的交叉之处，比如街心公园、道路的环状交叉口或中心广场的喷泉等。街区在我们的心理表征中占的空间很大，并以一些共同特征来命名，如剧院区或餐馆区等。最后，地标是在认知地图中可作为参照点的结构，我们通常在远处就能看到它，譬如钟楼、教堂的尖塔或者一座与众不同的高层建筑物等。

托尔曼在他早期的这篇论文中，清楚地阐述了认知地图理论，该文章自发表以来的50多年里，一直被众多研究领域里的研究者频繁引用。例如，最近有一项研究利用托尔曼的认知地图模型探讨了候鸟是如何依靠太阳的位置来确定路标，并形成惊人的成百乃至上千公里的迁徙认知地图（Bingman & Able，2002）。另一项来自旅游领域的研究也引用了托尔曼的思想，该研究的目的是检验那些到尚未开发的野外去旅游的游人（被称为自然旅游者）是怎样对那些地区的地形形成认识的（Young，1999）。该文的作者发现，有几个因素会影响被试心理地图的质量，其中包括交通方式、过去在当地的旅游经历、停留天数、旅游者的籍贯、旅游者的年龄以及性别等。

现在，许多旅行者根本不必外出，至少在身体层面上是这样。我们可以在互联网上到世界上任何一个地方旅游。托尔曼的认知地图理论也已在有关互联网的心理学研究中占有一席之地。想象一下你在互联网上所做的一切：你不断地在网上浏览，从一个地方跳转到另一个地方，在网上冲浪，四处游逛，用谷歌搜索。你常常感觉像在旅行，其实你的身体哪儿也没去。而且一旦有机会，你还会以几乎相同的路线去同样的地方，对吗？如果是，那么你已经对互联网中的某一小部分构建一幅心理地图！在《人—机关系》专刊中发表的一项研究就探寻了人们在互联网上的搜索行为以及人们在网上“冲浪”时所用的策略（Hodkinson et al.，2000）。研究者能够把网络搜索行为转化为图解形式，识别出个人的搜索策略，进而提出了一些提高网上搜索效率的可行方法。

将托尔曼的研究应用到一个古老的性别定型研究中，也许能够有所启发，即“男人永远不需要问路”。贝尔和索西亚（2004）探讨了人类的性别和性激素水平与他们沿着指定路线前进的能力之间的关系。设想一下，你正在沿着一条路从A处走向B处。走在路上，你会产生很多与周围环境相关的表象，比如一些明显的路标和沿途中令你感兴趣的地方。同时，在出发时，你会有一个整体的方向感。如果要求你指出其中的一些

心理表征，你可能会指认那些正确的位置，而不是其他的。换句话说，你已经建立起一幅属于你自己的认知地图，但它很难成为完美的。贝尔和索西亚要求这些被试完成类似的任务，结果发现性激素中最主要的睾丸素含量越高，越能显著地提高这些指向任务的准确性，这使我们较清晰地理解了被试在环境经历中形成的认知地图。那么，这是否意味着男人比女人问路少，是因为他们已经知道那个位置？并非如此。与这些结果同样吸引人的是，回答上述的问题还需要进行更多的研究。

Bell, S., & Saucier, D.(2004).Relationship among environmental pointing accuracy, mental rotation, sex, and hormones.Environment and Behavior, 36(2), 251-275.

Bingman, V., & Able, K.(2002).Maps in birds: Representational mechanisms and neural bases.Current Opinion in Neurobiology，12, 745-750.

Hodkinson, C., Kiel, G., & McColl-Kennedy, J.(2000).Consumer web search behavior: Diagrammatic illustration of way finding on the web.International Journal of Human-Computer Studies, 52(5), 805-830.

Lynch, K.(1960).The image of the city.Cambridge, MA: MIT Press.

Young, M.(1999).Cognitive maps of nature-based tourists.Annals of Tourism Research, 26(4), 817-839.

研究16 感谢记忆

Loftus, E.F.(1975).Leading questions and the eyewitness report.Cognitive Psychology, 7, 560-572.

佩里·梅森（Perry Mason）：汉密尔顿，我相信我的当事人所说的，她根本没有到过犯罪现场，我认为她的话是真实的。

汉密尔顿·伯格（Hamilton Burger）：佩里，我们为什么不让陪审团来裁定呢？

佩里·梅森：因为我认为这没有必要进行审判。你没有直接证据，你有的只是间接证据。

汉密尔顿·伯格：好吧，佩里，我想现在是告诉你一切的时候了。我们找到了目睹整个案件的人，佩里，我们有一个目击证人！

随着象征疑窦丛生的背景音乐逐渐响起，对佩里·梅森——这位最成功的律师而言，这又是一个非常棘手的案子。尽管我们有理由相信他将最终获胜，但对地方检察官而言，一个目击证人的出现完全可以使一个本来证据不足的案件一下子变得几乎无懈可击。为什么在刑事案件中目击证人的证词能提供如此确凿的证据？原因在于律师、法官、陪审团和普通民众都相信，一个人对一件事的记忆一定反映着该事件的本来面貌。换言之，记忆被认为是和录影机或DVD功能相似的回放事件的过程。然而，研究记忆的心理学家现在却对这种看法和其他有关记忆可靠性的观念产生了怀疑。

华盛顿大学的伊丽莎白·罗夫特斯（Elizabeth Loftus）是记忆研究领域中的卓越研究者之一。她发现，当人们回忆一件事情时，并不是准确地再现它；相反，回忆是一种对实际发生事件的一种重构。罗夫特斯的研究证明，重构记忆就是用新信息和现有信息去填补在回忆某种经历时被遗忘的内容。她认为人的记忆并不像我们通常认为的那么稳定，而是一段时间后就会被调整和改变。所以，如果你给别人讲述你五年前度假的事，你认为你是按照当时的情景来描述的，但事实可能并非如此。相反，你可能已经用了很多其他来源的信息重新建构了你的记忆，这些来源可以是你前几次对它的讲述、同一个或以后的假期中的其他经历，也可能是你去年看过的一部在你度假的地方拍摄的电影，等等。如果你和一个人在同时讲述你们共同的某种经历时，你经常会觉得奇怪，对同时经历的事件来说，你和另一个的叙述竟然会完全不同！

记忆的这些改变通常是无害的。然而，在司法程序中，当被告的命运可能由目击者的证词来决定的时候，记忆的重新构建就变得至关重要了。基于这个原因，罗夫特斯在记忆领域的很多研究都与法律上的目击者证词有关。在她的早期研究中，她发现，提问中措辞的细微差别都可

能改变一个人对事件的记忆。例如，如果问一个交通事故的目击者："你看到一个撞碎的车前灯了吗？"或"你看到那个撞碎的车前灯了吗？"尽管事件中根本就没有撞碎的车前灯，但在提问时使用"那个"比使用"一个"会产生更多的肯定反应。使用"那个"已经假定了事故中出现了撞碎的车前灯，这样就会使目击者在重构该事件时在原有的记忆中加入一个新的因素。

本节集中讨论的这篇文章是罗夫特斯最常被引用的几项研究之一，因为该论文分别报告了四项相关研究，而这四项研究把她的理论向前推进了一大步。她在这些研究中证明了，向目击证人提问时措辞的变化会干扰他们对事件的记忆，并影响他们对后面其他问题的回答。这项研究既影响了记忆理论，也影响了司法领域。

理论假设

该研究的焦点是探讨含有前提假定的问题对改变人们关于某个事件的记忆的作用。罗夫特斯把这种前提假定定义为：使问题很合理的、必须为真的一种条件。例如，假设你目击了一场交通事故，我问你："有几个人坐在这辆超速行驶的车里？"这个问题的前提假定为汽车是超速行驶的。但是如果这辆车实际上并没有超速行驶，你又该如何回答呢？你可能给出任何答案，因为这不是一个关于汽车是否超速的问题。然而，罗夫特斯认为，提问时的这种措辞方式会使你在这个事件的记忆中增加汽车超速行驶的这个信息。因此，在回答随后提出的其他问题时，你就更有可能说这辆汽车是超速行驶的。罗夫特斯假设，如果对目击者提出的问题中包含了关于目击事件的错误前提假定，那么这个新的错误信息可能会混入目击者对这个事件的记忆并且出现在目击者随后的新证词中。

方法和结果

以下是对四项实验中的每项实验的方法和结果部分进行的总结。

实验1

在第一项实验中，150名学生分别以小组为单位观看了一个短片，短片的内容为一个司机闯过了停车路标，冲进迎面而来的车流，并引发了五辆汽车的追尾事件。事件只持续了4秒钟，而整个电影也不超过1分钟。电影结束后，要求被试做一份包含10个问题的问卷。对于其中的一半被试，问卷中的第一个问题是这样问的："轿车A（闯过停车路标的那辆车）闯过停车路标时，速度有多快？"而另一半被试的第一个问题是："轿车A右转弯时的速度有多快？"接下来的一些问题对研究者而言

都是无关紧要的，他们真正感兴趣的是最后一个问题。该问题对两组被试的提法完全相同，即“轿车A前面是否有停车路标？”

对在第一个问题中提到停车路标的那组被试，有40名（53%）说他们看到轿车A前有停车路标，但在“右转弯”组中，只有26名被试（35%）声称他们看到了停车路标。这种差异在统计上达到显著。

实验2

罗夫特斯的第二项研究是在她的系列研究中首个包含延迟记忆测试的研究，也是四项实验中唯一一项没有使用交通事故作为目击事件的研究。在这项研究中，主试给40名被试放一段3分钟的电影片段，片名为《学生革命日记》。这段电影片段说的是一个班级遭到了8名反战示威者的捣乱。被试看完电影后，要求他们回答问卷，该问卷含20道有关该电影片段的问题。其中一半被试的一个问题是：“进入教室的4名示威者的带头者是男性吗？”另一半被试的问题则是：“进入教室的12名示威者的带头者是男性吗？”向两组被试提问的其他问题完全相同。

在初测的一个星期之后，主试把两组被试请回并请他们回答有关这部电影的20道新问题（不再重看电影）。与研究结果相关的一个问题是：“你看见几名示威者走进了教室？”别忘了，两组被试看的是同一部电影，而且所有问题中除了作为参考的12名示威者与4名示威者的不同外，其余全部相同。

在问题中假定了12名示威者的那组被试平均报告的人数为8.85。在问题中假定了4名示威者的那组被试平均报告的人数为6.40，这也同样具有显著性差异。这个实验表明，平均来看，问题的措辞改变了被试对目击事件基本特征的记忆。

实验3

研究者设计这个实验的目的是想知道，提问中的一些错误假定是否会使目击者在重建对某事件的记忆时把未曾出现在事件中的物体包括进来。被试（150名大学生）观看一段与一辆白色赛车有关的交通事故的短片，录像的内容围绕着这一事故展开，而后要求他们回答与录像有关的10个问题。一半被试的问卷中包含的一个问题是“白色赛车在乡间道路上行驶，它经过谷仓时速度有多快？”对另一半被试则问道：“白色赛车在乡间道路上行驶时，速度有多快？”如同前面的研究一样，被试一个星期后返回并回答10道与这个事件有关的新问题。其中，用来验证研究假设的问题是：“你是否看见了一个谷仓？”

在前一轮问题中提到谷仓的那组被试中，有13人（17.3%）对此问题回答“是”，而相比之下，没有提到谷仓的那组被试中只有2人（2.7%）回答“是”。两组的差异又一次达到了统计上的显著性水平。

实验4

该论文的最后一项实验是为了实现两个目标而精心设计的：第一，罗夫特斯希望进一步证明在实验3中发现的记忆重构效应；第二，她想知道，仅仅提及某物，即使它不是错误前提假定的一部分，是否也足以使被试把它补充到记忆中去。例如，当影片中没有出现谷仓，主试直接问："你是否看到一个谷仓？"你可能回答："没有。"但如果主试一周以后再问你同样的问题，谷仓就可能会潜入你对该事件的记忆里。这就是罗夫特斯要在她的第四项实验中验证的想法。

研究者将50名被试分为3组，让他们观看了一段三分钟的短片，电影是在一辆汽车内拍摄的，结尾是这辆车撞到了一位男士推着的婴儿车上。然后向3组被试分发包含有关该短片内容的问题册子。这些小册子有如下不同：

D组：直接提问组。他们收到的小册子中包含了40道填充题和5道关键的直接提问的题，这5道题均包含了事件中并不存在的事物。如，"你是否在电影中看到一个谷仓？"

F组：错误前提假定组。他们收到了同样的40道填充题和5道关键问题，这些关键问题包含了对不存在物体（与D组相同）的前提假定。如，"你是否看见有一辆旅游车停放在谷仓前？"

C组：控制组。他们只需回答40道填充题。

一周后，所有被试返回并回答有关该短片的20个新问题。这些问题中有5道与一周前的直接提问组回答的关键问题是完全相同的。所以，D组对这5个问题回答了两遍。因变量（结果）是计算每组中报告自己记得在短片中出现了某事物而实际并不存在的人数百分比。

表4—2展示了三组被试的回答结果。请记住，短片中并没有校车、卡车、道路中央线、推着婴儿车的女士以及谷仓。把所有问题综合在一起，一周后对这5道题回答"是"的总百分比具体如下：错误前提假定组为29.2%，直接提问组为15.6%，控制组为8.4%。直接提问组与错误前提假定组在每一问题上的差异，以及所有问题的综合差异都达到统计上的显著性。

表4—2 经过直接提问和错误假设前提提问之后，被试对电影中并不存在的物体的回忆次数

直接提问	错误的前提假定	一周后对直接问题回答为“是”的百分比		
		C	D	F
在电影中看到校车了吗?	看到孩子们上校车了吗?	6	12	26
电影开始时看到卡车了吗?	电影开始时，卡车是停在汽车边上的吗?	0	8	22
在乡间道路上看到中央线了吗?	是否有另一辆汽车越过了乡间道路的中央线?	8	14	26
看到一位推着婴儿车的女士吗?	该女士是否推着婴儿车穿过道路?	26	36	54
在电影中看到谷仓了吗?	看到一辆旅行车停在谷仓之前吗?	2	8	18

C：控制组　　D：直接提问组　　F：错误前提假定组

（引自 P. 568）

讨论

在上述研究及其他研究的基础上，罗夫特斯主张一种有关记忆和回忆的准确理论必须包括重构过程，即新的信息被整合到人对某一事件的最初记忆中去。这些研究结果并不能用回忆仅仅是心理上对一个事件的重放这一理论假设来解释，即使这种重放可以有不同的准确度。为了说明这些问题，图4—8对传统的回忆观点与罗夫特斯提出的新观点作了比较。新观点中增加了新信息整合到记忆中的过程，而传统观点中并不包含这一点。进而，这种新信息会使你原先的记忆表征得以改变或重构。随后，如果问你一个有关这个事件的问题，你回忆的将不再是原先实际发生过的事件，而是你对它的重构。罗夫特斯认为，重构过程正是谷仓、校车、推着婴儿车的女士和道路中央线这些在短片中根本不存在的事物会鬼使神差地出现在被试记忆中的原因。在提问中错误的前提假定为人们提供的新信息被无意识地整合进了最初的记忆中。

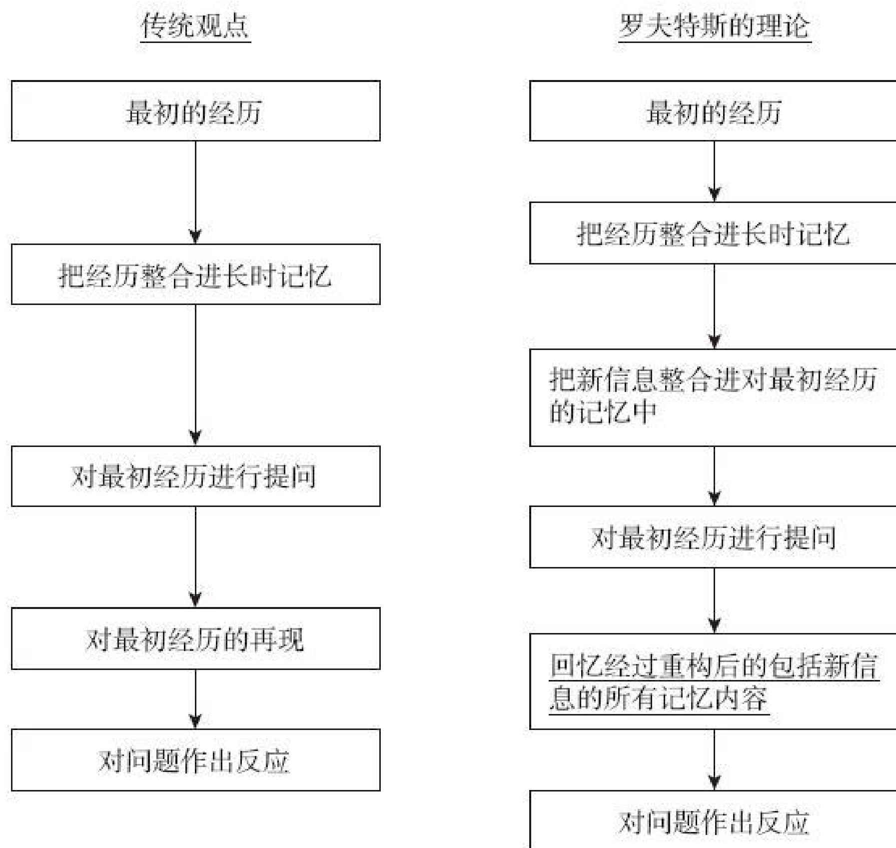


图4—8 回答问题时对事件的回忆

把该理论应用到刑事调查中的目击者证词上时，罗夫特斯指出，在刑事调查中目击证人通常要多次回答问题。他们可能在犯罪现场接受警察的提问、接受负责处理该案的检察官的询问，而且出庭作证时将再次被询问。在上述不同时间段的提问过程中，很有可能会产生错误的前提假定，而且会以许多种方式出现，尽管这可能并非是有意的。研究表明，一些常见的、听起来没有问题的提问，如“那家伙的枪是什么样的？”或“逃逸的汽车停在哪儿？”等，无论枪或逃逸的汽车是否真正存在，这样的提问都会提高目击者记住它们的概率（Smith & Ellsworth，1987）。所以，虽然目击证人、检察官、法官和陪审团都认为，目击证人是在再现他们的所见所闻，但罗夫特斯认为，目击证人记住的是“建立在改变了的记忆表征基础上的再生表象”（P.571）。

近期应用

罗夫特斯在目击证人证词方面所做的一系列令人印象深刻的研究对相关研究产生了巨大影响，以下几项研究就是很好的例证。其中一项研

究引用了罗夫特斯这篇1975年发表的论文，该研究考察了律师的复杂提问方式如何对目击证人证词的准确性和目击证人的信心产生负面影响（Kebbell & Giles，2000）。所有被试观看同一段录像，一周后主试对他们所看到的内容进行提问。对一半被试提问时使用的是含糊不清的语言（律师们常会说："这……并不真实，对吗？"），而对另一半被试提问时，则使用相同问题的简练表达方式。结果很清楚：回答含糊不清的问题的被试与回答直截了当的问题的被试相比，他们所报告的证词不太准确，而且对答案也缺乏信心。其他研究证实了当有多个嫌疑犯同时出现在目击证人前时（这在执法过程中是非常常见的），目击证人辨认嫌疑犯的准确度显著下降，因为他们将这些新面孔纳入到了对原始案件的重构过程（Pezdek & Blandon-Gitlin，2005）。

另一项有趣的研究将罗夫特斯的成果应用到"幻想记忆"（fantastic memories）中，即相对于现实情况来说，更像是幻想出来的记忆，如被外星人绑架、灵魂出窍经历、超感觉（ESP）事件、遭遇鬼魂，等等（French，2003）。很明显，如果有关这些记忆的报告是准确的，则将会为这些"异常"事件的发生提供证据。然而，研究结果反复地告诉我们这样的事件从未得到科学证实。那么，如何对这些记忆作出解释呢？答案可能就是我们本章所讨论的人类记忆的易误性和不可靠性，而且还可能是我们大脑有能力创造出对实际上从未发生过的事件的记忆。正如弗兰奇（French）指出"大量的心理变量被证明与错误记忆的易感性有关（例如催眠易感性、分裂倾向），同时也和报告异常体验有关"（French，P.153）。

伊丽莎白·罗夫特斯除了研究目击者证词领域之外，她现在还是研究被压抑的童年期记忆的代表人物，而这一问题则引起了颇多争议。其中争论的一方声称自己在过去的生活中曾经被虐待，且通常是性虐待，但他们只是在最近的心理治疗师的帮助下，才记起了自己的受虐经历，因为这些创伤性记忆已经被压抑到他们的潜意识中去了。另一方面的争论是关于那些被谴责为施虐者的人，他们坚决否认，并认为这些记忆既可能是幻觉，也可能是通过治疗过程灌输给他们的（参见Garry & Loftus，1994，对该争论的回顾）。这正好属于罗夫特斯记忆研究的范畴。

罗夫特斯在其著作《被压抑的记忆之谜：错误记忆和对性虐待的辩解》（Loftus & Ketcham，1994）中总结了她在该领域的研究结果，并把它们综合成了一个连贯的论证观点。从本质上说，罗夫特斯主张根本不存在被压抑的记忆，而且她在很多研究中似乎已经证明了这一点。事实上，她是质疑潜意识的整个概念及其真实存在的最前沿的心理学家。罗夫特斯观点中的一个主要特征是，实验证据反复证明了创伤性记忆往往是我们记得最清楚的部分。然而，临床心理医生却报告说，这些被压抑的性虐待记忆是在特定的深入治疗过程中逐渐浮现出来的，即从无意识状态进入意识状态。如何把这两个看似对立的观点统一起来呢？罗夫特斯提出三种可能的记忆歪曲过程用来解释被临床心理医生称为"压抑"的现象（Loftus，Joslyn，& Polage，1998）。首先，当事人也许早

已忘记了早期的性虐待，而不是压抑了它们。她引用的研究证据是，如果儿童根本不理解虐待事件中性的本质时，那么虐待事件往往很难被儿童记住。其次，人们可能在治疗中说他们已经忘记了某一创伤性事件，但事实上他们根本没有忘记。不愿想起某件事并不等于忘记了此事。最后，罗夫特斯坚决主张，“某些人可能相信某种特定的创伤性事件的确发生过而且被压抑了，但实际上，这件事最初并没有发生过。只是在某种情况下，这些歪曲了的记忆结合在一起导致了‘压抑’的出现”（P.781）。

不难想象，罗夫特斯关于被压抑的和被恢复的记忆的见解并非没有招致批评（参见Spitzer & Avis，2006；Steinberg，2000）。毕竟，她对压抑力量的否认直接否定了弗洛伊德提出的心理学和心理治疗的观点。另外，很多心理治疗师和病人从个人角度出发，都相信与受虐待有关的记忆可被压抑很多年，而后得以恢复。然而，仔细阅读罗夫特斯全面而细致的科学研究著作应该会使每个人都对这一观念提出质疑。

结论

伊丽莎白·罗夫特斯被公认为记忆重构及目击者证词研究领域的卓越先导。她在这方面的研究还在继续进行。多年来，她的研究结果成功地经受住了各种挑战，并且得到了该领域其他研究者的支持。

在当今的心理学和法学专业领域内，人们普遍相信在目击者的证词中会掺杂很多来自不同渠道的错误信息，例如混入一些该事件发生后的信息。正是由于罗夫特斯和其他研究者的一系列研究，目击者的证词在审判中的作用和可靠性才受到了严重质疑。罗夫特斯本人就是最为人们器重的作证专家之一（通常是为被告作辩护）。她的职责是用事实来提醒陪审团成员，在评估目击者的证词时必须保持谨慎。

正像罗夫特斯在她最近的著作中写到的：“我研究记忆，同时我是一个怀疑论者。”（Loftus & Ketcham，1994，P.7）或许，我们都应该如此。

French, C.(2003).Fantastic memories: “The relevance of research into eyewitness testimony and false memories for reports of anomalous experiences.Journal of Consciousness Studies，10，153-174.

Garry, M., & Loftus, E.(1994).Repressed memories of childhood trauma: Could some of them be suggested? USA Today magazine，122, 82-85.

Kebbell, M., & Giles, C.(2000).Some experimental influences of

lawyers' complicated questions on eyewitness confidence and accuracy.Journal of Psychology，134(2)，129-139.

Loftus, E., Joslyn, S., & Polage, D.(1998).Repression: A mistaken impression? Development and Psychopathology，10(4), 781-792.

Loftus, E., & Ketcham, K.(1994).The myth of repressed memories: False accusations and allegations of sexual abuse.New York: St.Martin's Press.

Pezdek, K., & Blandon-Gitlin, I.(2005).When is an intervening line-up most likely to affect eyewitness identification accuracy? Legal and Criminological Psychology，10(2), 247-263.

Smith, V., & Ellsworth, P.(1987).The social psychology of eyewitness accuracy: Leading questions and communicator expertise.Journal of Applied Psychology, 72, 294-300.

Spitzer, B., & Avis, J.M.(2006).Recounting graphic sexual abuse memories in therapy: The impact on women's healing.Journal of family Violence, 21(3)，173-184.

Steinberg, M.(2000).The stranger in the mirror.Psychology Today, 33, 34.

第五章 人的发展

心理学中关于人的发展这一分支关注每个人从出生到死亡的整个过程的发展变化，这是行为科学研究中最庞大、最复杂的课题之一。虽然每个人都是独一无二的个体，但我们的成长过程是相似且能被预测的，而且有一个相对固定"进程表"。发展心理学中最有影响的研究领域是婴儿与母亲之间的依恋过程，智力的发展和年龄老长带来的变化。

在这一部分中，我们要讨论一些在心理学领域中最著名和最具影响的研究。亨利·哈洛（Harry Harlow）博士用猴子做的实验证明了婴儿早期的依恋对以后的心理调节的重要性。让·皮亚杰（Jean Piaget）包罗万象的研究成果，为今天人尽皆知的认知发展理论奠定了坚实的基础。在此，我们将详细介绍他的一项研究，使读者得以管窥其巧妙的研究方法及清晰的研究结果。另外，我们还将介绍劳伦斯·柯尔伯格（Lawrence Kohlberg）的著名研究，其内容是关于人类的道德品质是如何发展的，以及为什么一些人表现出更高的道德水平。此外，由于人的发展是个终其一生的过程，我们还将探讨由艾伦·兰格（Ellen Langer）和朱迪斯·罗丁（Judith Rodin）所做的著名研究（通常也叫"植物研究"），以此来说明任何人无论他的年龄多大，都需要拥有能控制自己选择、行为和命运的能力。

研究17 爱的发现

Harlow, H.F.(1958).The nature of love.American Psychologist，13, 673-685.

有时候，心理学家似乎是走得太远了。像“爱”这类问题怎么可能用科学的方法进行研究呢？但无论你怎样界定“爱”的含义，你不得不承认它对我们的行为有巨大的影响。如果我们作出一个假设，那么接下来心理学家就会提出一系列他们所感兴趣的东西，比如什么是爱，从哪里会得到爱，爱是如何起作用的。

发展心理学家亨利·哈洛（Henry Harlow，1906-1981）被公认为自弗洛伊德之后在研究早期经验对成年的影响方面作出巨大贡献的心理学家。绝大多数心理学家都同意哈洛的观点：婴儿与母亲（或者早期看护者）之间的亲密接触和依恋经历对其在今后生活中爱的能力及与他人亲近的能力有很重要的影响。仔细想想，在你生命中，最早的有关爱的经历是什么？那就是自你出生伊始便存在的与母亲之间的联系。那么，这种接触确实非常重要吗？弗洛伊德主义者相信，在生命开始的第一年里，亲密接触主要集中在乳房和本能的口唇需要（著名的口唇期）。后来，行为主义者反对这种观念，并认为人的所有行为都与本能需要（如饥饿、干渴和回避痛苦等）有关。由于母亲可以满足这种需求，所以婴儿与母亲之间的亲密关系会在母亲喂养婴儿的过程中不断得到强化。因而，母亲就与愉快的事件联系在一起，于是爱就产生了。在这两种观点中，爱都是其他本能或生存需要的附属品。然而，亨利·哈洛却发现，与饥饿和干渴一样，爱和情感也是一种强烈的基本内在需要，甚至比前者更强烈。

用以揭示婴儿与母亲间爱的成分的方法之一，就是把婴儿放在一种特殊环境里，在这种环境中，母亲不再满足婴儿的任何需求，同时研究者可对这种环境中的多种成分进行科学的操纵。按照上述理论，我们通过改变母亲满足婴儿基本需要的能力，便能阻止或改变婴儿与母亲之间所形成的具有一定质量和强度的依恋。然而，由于伦理道德的原因，这种实验很显然是不能在人类身上实施的。由于哈洛用恒河猴进行学习课题的研究已有好几年了，所以他便用猴子作为被试，进行他的爱与依恋的实验研究。从生理学角度看，恒河猴与人类非常接近。哈洛也相信，恒河猴在婴儿期对情感和接触（如喂养、接触、依附等）的基本反应与人类相同。那么，用动物作被试进行这种研究是否存在伦理问题呢？我们将稍后对此进行讨论。

理论假设

在哈洛的早期研究中，幼猴在实验室接受精心的人工抚养，研究者用奶瓶悉心喂养它们，挑选食物使它们吸收均衡的营养，并使其免受疾病威胁。在这种情况下成长的幼猴比由母猴照顾的幼猴更健康。哈洛注意到这些幼猴非常依恋盖在笼子底部的布垫子（棉花垫子）。当研究者把这些垫子拿去清洗时，它们变得非常生气和焦虑。他还发现仅出生一天的幼猴便能表现出这种依恋，其程度在出生后的最初几个月中会变得越发强烈。很明显，就像哈洛所说的那样："对于婴儿而言，无论是婴儿，还是幼猴，为了生存，他们必须抓住比稻草更多一些的东西。"（P.675）如果幼猴生活在一个没有软垫覆盖的笼子里，即使它拥有非常好的营养和医疗条件，它也无法茁壮成长。把布垫放进去后，幼猴就变得更健康，看起来也更为满足快活。所以，哈洛的理论认为，幼猴除了基本的饥饿、干渴等生理需求外，它们一定还有一种要接触柔软物质的需求。为了验证这个理论，哈洛和他的合作者决定"制作"各种类型的代理母猴。

方法

他们制作的第一只代理母猴是这样的：用光滑的木头做身子，用海绵和毛织物把它裹起来；在胸前安装了一个奶瓶，身体内还安装了一个提供温暖的灯泡。然后他们又组装了另一只不能提供舒适环境的代理母猴。这只母猴是由铁丝网制成，外形与绒布母猴基本相同，以便使幼猴用接近木猴的方式接近它。这只铁丝母猴也安装了能喂奶的"乳房"，且也能提供热量。换言之，这只铁丝母猴与绒布母猴相比，除了在被哈洛称为"接触安抚"的能力方面有差异外，其他方面完全一样。

然后，研究者把这些人造母猴分别放在单独的房间里，这些房间与幼猴的笼子相通。八只幼猴被随机分成两组，一组由绒布母猴喂养（用奶瓶），另外一组由铁丝母猴供奶喂养。我相信，你已经很清楚哈洛为什么要做这个实验。他企图将喂养的作用与接触安抚的作用分离开来。哈洛把猴子放在笼子里，并记下在出生后的头五个月中，幼猴与两位"母亲"直接接触的时间总量。结果是令人惊讶的，我们将稍后对此进行讨论。

在完成了这些最初的研究后，哈洛想进一步探索"依恋"及接触安慰的作用。一般的常识告诉我们，当孩子们感到害怕时，他们总会到母亲（或者其他早期看护者）那里寻找庇护。为了探寻在这种情境下，与铁丝母猴在一起的幼猴和与绒布母猴在一起的幼猴将分别作出何种反应，哈洛在它们的笼子里放入各种各样能引发恐惧的物品，如上紧发条的玩具打鼓熊（这种玩具熊与幼猴一样大，对幼猴而言是很可怕的）。研究者对在这些情况下幼猴的反应进行观察，并作了详细记录。

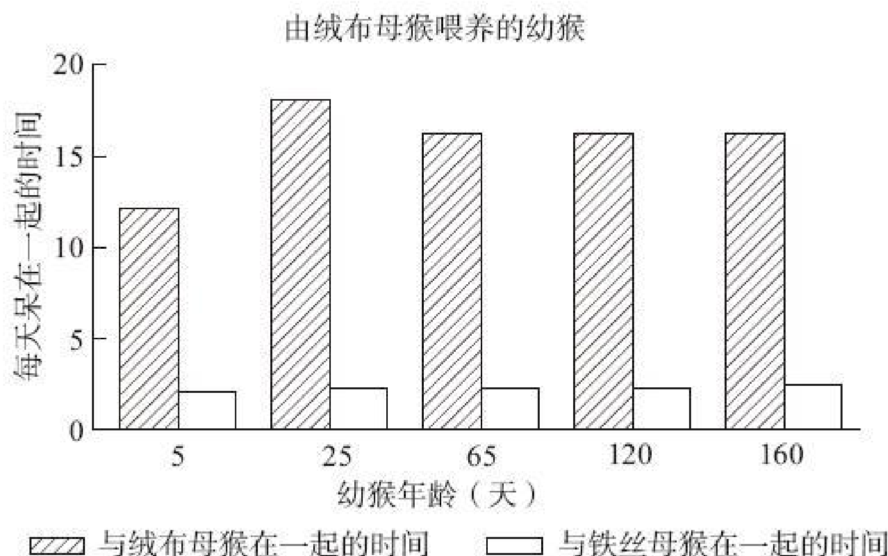
哈洛的另一项研究被称为"旷场试验"，他把幼猴放进一个不熟悉的小房间里，里边放着各种各样的物品（如积木、毯子、带盖的容器、折

纸等）。在一般情况下，猴子喜欢玩这些玩具或摆弄它们。把两组幼猴放进同一个房间，并设置三种情况：仅出现绒布母猴；仅出现铁丝母猴；两者都不出现。哈洛的想法是要考察母猴在场或不在场的情况下，这些幼猴适应和探索这种陌生环境的倾向性。

最后，哈洛想探索幼猴与代理母猴之间形成的依恋关系是否在它们分开一段时间后还能保持。当幼猴长到六个月能够吃固体食物时，让它们与“母亲”分开一段时间，然后再在旷场环境中团聚。

结果

你可能还记得，在最初的实验中，所有的幼猴与两只代理母猴都接触。其中一半幼猴由绒布母猴喂奶，另一半则由铁丝母猴喂奶。现在，你可能已猜到幼猴偏爱的是由绒布包裹的木制母猴（不是吗？），但是令人惊奇的是，这种偏爱程度趋向于极端，甚至对那些由铁丝母猴喂养的幼猴而言也是如此。母猴是否满足幼猴的饥饿、干渴等生理需求并不是幼猴依恋母猴的主要因素，这与当时流行的观点恰恰相反。接触安抚在幼猴对母猴产生依恋的过程中有重要影响，这一点在实验中得到了清楚的证明。图5—1用图表的方式说明了这一结果。经过最初几天的调适后，无论哪只母猴提供奶，所有的幼猴几乎整天与绒布母猴待在一起。甚至是那些由铁丝母猴喂养的幼猴，它们为了吃奶才迫不得已离开绒布母猴，吃完后便迅速地返回到绒布母猴这里。



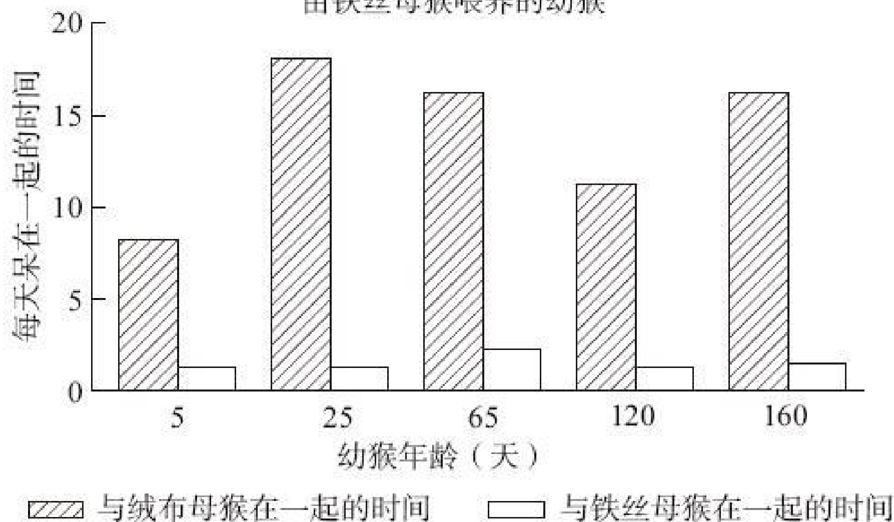


图5—1 幼猴每天与绒布母猴和铁丝母猴待在一起的时间

分别由绒布母猴和铁丝母猴喂养的两组猴子的行为特征进一步证明了接触安抚的重要性。虽然两组猴子食量同样大，体重增长的速度也基本相同，但由铁丝母猴喂养的幼猴对牛奶消化不良，且经常腹泻。这意味着，缺少母亲的接触安抚使幼猴产生了心理上的紧张。

恐惧物体的实验结果进一步证明了幼猴对绒布母猴的依恋。每当幼猴发现自己正面对一些害怕的事物时，它们便很快跑向绒布母猴，并抱住它以获得安慰和保护。随着幼猴年龄的增长，这种反应变得愈发强烈。另外，无论是铁丝母猴喂养的幼猴，还是绒布母猴喂养的幼猴，其反应都没有差异：当它们害怕时，都会到绒布包裹的代理母猴那里寻求安全感。

你也许曾注意到，当父亲或母亲在场时，孩子们会感到安全与放心，他们在这种场合下更充满好奇心，更乐意去探索他们周围的环境。通常，他们会探究周围环境中的一切，但其前提是他们必须看到父母在场。哈洛的陌生环境或称作旷场试验就是这样设计的，以此来引发猴子产生类似的行为。研究者把幼猴放入陌生的环境后，所有幼猴立即冲向绒布母猴，抓住它，用身体蹭它，并摆弄它的脸和身体。一会儿以后，这些幼猴“开始把绒布母猴看作安全之源……它们在这个陌生的新环境里探索和摆弄各种物品，然后返回到母亲怀里，循环往复”（P.679）。

然而，如果我们把这些幼猴放在同一间房间里，但绒布母猴不出现时，它们的反应就完全不同了。他们充满了恐惧，出现情绪化的行为，如哭叫、缩成一团、吸吮手指。有时候，它们会跑向房间里母猴曾出现过的那个地方，然后从一个物体跑向另一个物体，尖声嘶叫着。在铁丝母猴出现的情况下，幼猴的行为表现与它们在两种母猴都不出现的情况下的表现是完全一样的。所有幼猴都是这样，无论它们是由谁（绒布母

猴或铁丝母猴）喂养的。

在该研究的最后一部分中，研究者让已过哺乳期、可食用固体食物的幼猴（大概5~6个月大）与母猴分离（最长时间约30天）。当幼猴们在相同的旷场环境中再一次与绒布母猴重逢时，它们冲向母猴，爬在它身上，紧紧抓住它，用自己的头和脸在它身上磨擦，然后与母猴玩耍，撕咬包裹在母猴身上的绒布。最明显的变化是幼猴不再像以前那样，离开母猴去探索和玩耍房间里的其他物品。按照哈洛的观点，很显然这说明寻找安全感的需要比探索环境的趋向更为强烈。然而，需要指出的是，重聚的时间仅仅持续了大约3分钟，假如我们把这段时间再延长一些，可能会产生更多的探索行为。

讨论

哈洛指出，他的研究证明，接触安抚对幼猴与母猴间依恋关系的发展具有极其重要的作用。事实上，对于幼猴而言，接触安慰在依恋关系的形成中比母猴提供乳汁的能力更重要。

这项研究改变了心理学的原因之一在于，该研究所得出的结论与当时所流行的行为主义观点背道而驰。行为主义的观点认为，推动母婴关系发展的动力是喂养行为所导致的强化。然而，正如哈洛所说："作为一个感情变量，早期哺乳行为使婴儿与母亲之间经常发生亲密的身体接触。当然，人是不能仅仅依靠乳汁来生活的。"（P.677）

毫无疑问，哈洛（和许多其他研究者）认为他的研究结果应该被推广到人类，我们稍后将很快会谈到这个问题。实际上，他已为研究成果应用于人类提供了可能性。哈洛认为，随着社会经济对家庭的要求不断增长，会有越来越多的妇女进入社会，参加工作。这也是那个年代许多人对哈洛的研究给予关注的原因之一。因为，那时人们仍普遍认为，母亲亲自喂养孩子对孩子的健康成长和情感发展来说是必不可少的。他说，成功养育的关键是接触安抚，而不仅仅是妇女的哺乳能力，美国男性也能在养育婴儿方面起到相同的作用。今天，这种观点已为人们广泛接受，但是，哈洛在1958年写这篇报告时，这简直如同一场革命。

批评与研究发展的意义

尽管哈洛坚持这个主张，但是你是否也认为人类与猴子的依恋或爱的过程是相同的？有一些研究支持这个观点，即婴儿对其照看者的依恋确实超越了生理需求的满足。我们可以很清楚地看到，母亲与她的孩子之间的肌肤接触增进了依恋关系（Klaus & Kennell，1976）。但是，对于人类而言，这种依恋的发展过程要缓慢得多：六个月大的人类婴儿的

依恋水平大约只相当于刚出生几天的幼猴。另外，大约仅仅有70%的孩子在一岁的时候才显现出与成人之间的安全依恋关系（Sroufe，1985）。

无论过去还是现在，有许多人基于伦理原则对哈洛用幼猴做实验的研究工作提出了批评。他们提出了这样的问题：作为人类，我们有权在有潜在危害的情境中用猴子或其他动物做实验吗？就哈洛的研究而言，在这个问题上的争论，双方都有合情合理的解释。科学界判断此项研究是否符合伦理原则的途径之一是看其是否对人类和社会有潜在的益处。无论你认为这项研究是否符合道德标准，它已经对人类社会的许多方面产生了积极的影响，有时还涉及收养儿童或虐待儿童等问题。

不幸的是，在我们的文化背景下，有许多孩子被迫长时间地生活在福利院等机构之中，有些是因为他们的父母不能够照顾他们（孤儿），有些是因为他们自身有疾病或是残疾。哈洛的研究已经影响了我们对这些孩子的照顾方式，现在人们普遍认为，仅满足这些孩子的生理需求是远远不够的，我们应使孩子有更多机会与他人进行身体接触。所以，工作人员、护士、志愿者等应尽可能多地抚摸和拥抱这些孩子。而在医疗条件有所限制的情况下，通常我们把这些孩子放在他们彼此能够看见、相互能触摸的环境中，以使他们获得更多一些的接触安抚。虽然在满足依恋需要方面，这种努力将永远不能真正替代来自父母的关怀，但这与传统的看护方式相比，着实前进了一大步。

哈洛的工作还鼓舞了意欲成为好家长的男性看护者，使他们信心倍增。在婴儿的发展与适应过程中，接触安抚与哺乳相比更为重要，因此孩子的生母不再被看作是照顾婴儿的唯一合适人选。今天，父亲完全可以在这个过程中承担起一个重要的角色。除此之外，在必要时其他除父母外的看护者，如保姆或日托中心的工作人员，都可成为婴儿合适的看护者。而且，这些发现大大扩展了收养子女的前景。因为人们已认识到，养父母与亲生父母一样，可以给孩子提供同样的接触安抚。

哈洛的早期研究最后对虐待儿童问题进行了一些解释。令人惊讶的是，几乎在所有案例中，那些被虐待的孩子似乎很爱甚至依恋那些对其施虐的父母。对严格的行为主义者而言，这是难以理解的。哈洛解释说，如果依恋是最强烈的基本需求，那么它的作用将远远超出施虐惩罚。实际上，哈洛在后来的实验中检验过这一观点。他设计了能够拒绝幼猴的代理母猴，即让一些母猴向空气中喷射强气流，与此同时，其他母猴用钝头针突袭幼猴，迫使幼猴离它而去。幼猴对这种情形的反应是，与母猴保持一小段距离，等这种驱逐行为一结束，它们便再返回来，与先前一样紧紧地抓住母猴（Rosenblum & Harlow，1963）。

近期应用

哈洛的研究不断被其他研究广泛引用，这些研究包括探讨接触、亲密、依恋的影响以及人与人之间的交际对情绪和生理健康的影响作用。其中有一篇研究探索了那些处于孤独生活情境下的成年人的社会隔离（缺乏与他人建立亲密的、有意义的社会联系）和生理健康之间的关系（Cacioppo & Hawley，2003）。研究结果表明，与那些拥有健康社会交际的人们相比，缺乏交际的人会觉得日常的生活事件压力更大，患上高血压的风险更高，在受伤后治疗恢复得更缓慢，睡眠也较为不好。

另外一项引用哈洛研究成果的研究已证明，母婴之间的肌肤接触（有人将这形象地比喻为“袋鼠式的照顾”）对早产婴儿的存活和发展以及母婴之间亲密关系的建立是至关重要的（Feldman & Eidelman，1998）。这是一个非常重要的发现，在护理高危早产婴儿的医院里，医生、护士既要满足婴儿身体接触的需要，又要避免那些潜在的、威胁婴儿生命的传染病，因为婴儿尚未发育完全的免疫系统不具有抵抗力。

哈洛的研究结果还适用于心理治疗机构。在过去的40年中，人本主义和整体疗法已有了长足的进步，触摸治疗在这一领域扮演着越来越重要的角色（参见LaTorre，2000）。就像一位心理治疗师所述：

> 我已经发现，尤其是在穴位按摩和治疗性接触之类的接触疗法中，专注及有意识的触摸会成为治疗互动的一个重要部分。它能深化治疗中的认识并可对治疗效果提供帮助。触摸疗法若用得其所，可以增强而绝非削弱心理治疗的互动。但要注意的是，这里的关键词正是“用得其所”。接触是强有效的工具，要多加利用（LaTorre，2000，P.105）。

结论

如果有人认为哈洛对爱的定义有独断性，那将是错误的。正是他的发现改变了我们对于母婴之间关系的看法，这一点却是确定无疑的。如果这个研究已经渗入到我们的文化当中，哪怕只是一点点，我们也许就能看到一些好的结果。例如，哈洛曾提到一名妇女，这名妇女听完了他的研究报告后，来到他面前说：“现在我知道我的问题到底出在哪儿了，我就像是那只铁丝母猴。”（P.677）

Cacioppo, J., & Hawkley, L.(2003).Social isolation and health with an emphasis on underlying mechanisms.Perspectives in Biology and Medicine, 46, S39-S52.

Feldman, R., & Eidelman, A.(1998).Intervention programs for premature infants: How and do they affect development? Clinics in Perinatology, 25(3), 613-629.

Klaus, M.H., & Kennell, J.H.(1976).Maternal infant bonding.St.Louis: Mosby Press.

LaTorre M.(2000).Integrative perspectives.Touch and psychotherapy.Perspectives in Psychiatric Care.36 ，105-106.

Rosenblum, L.A., & Harlow, H.(1963).Approach-avoidance conflict in the mother surrogate situation.Psychological Reports ，12, 83-85.

Sroufe, A.(1985).Attachment classification from the perspective of the infant-caregiver relationships and infant temperament.Child Development, 56 ，1-14.

研究18 眼不见，不一定心不烦

Piaget, J.(1954).The development of object concept: The construction of reality in the child (pp.3-96).New York: Basic Books.

你是如何从一个仅有基本思维能力的婴儿发展成现在能使用语言、符号和逻辑等众多复杂方法对事物进行分析、推理的成人的？对于这个问题，你的第一反应很可能是这样的：“嗯，我是通过学习获得这些智能的。这是一个我与外界环境相互作用的过程，也是从成人那里接受教育的过程。”

仅凭直觉，大多数人对于这个解释会持肯定态度，但许多发展心理学家却认为，智能的获得不仅仅只是通过学习，还有其他许多渠道。人们普遍认为，智力的发展是一个成熟的过程，就像身体的成熟一样，从出生到长大成人，它始终是以可预测的方式进行的。

当你面对一个婴儿时，你会认为只要通过足够的学习，他便会表现出许多成人的行为吗？当然不会。相反，你明白，这里有一个发育成熟的过程，它将使孩子在未来表现出越发复杂的行为。只有当孩子们达到了特定的发展水平，某种学习才能导致某种行为。例如，以走路这个行为来说，你也许会认为走路是一种习得的行为。但设想一下，教一个6个月大的婴儿走路，即使你按奥林匹克训练计划每天对婴儿进行8小时的训练，他也学不会走路，这是因为他的身体还没有成熟到能够走路的程度。

大多数研究者认为，智力或认知的发展在很大程度上也是这样一个过程。当认知发展达到某个特定的阶段，个体才会有特定的推理和思维水平。否则，无论进行多少次学习，也学不会这种能力。心理学界把对认知发展概念的这种理解与认识归功于瑞士心理学家让·皮亚杰（Jean Piaget，1896～1980）。

让·皮亚杰是心理学史上最有影响的人物之一，他的研究不仅引发了发展心理学的一场革命，并为其后的智力结构研究奠定了基础。让·皮亚杰早先攻读生物学，主要致力于研究动物适应环境的先天能力。在巴黎索邦神学院（即巴黎大学的前身）研究期间，他在阿尔弗雷德·比纳（Alfred Binet）实验室找了一份工作，以赚取额外报酬，这个实验室便是世界上第一个智力测验的诞生地。他的工作是将一项用英语开发的推理测试的法语版标准化。正是在巴黎工作的这段时间里，让·皮亚杰开始形成了他自己的认知发展理论。

理论假设

起初，对于皮亚杰来说，比纳实验室的工作是枯燥乏味的。但在随后的测验中，不同年龄段的儿童在回答问题时表现出的一些有趣模式开始引起他的注意，即同一个年龄段的儿童出现的错误是相同的，也就是说，他们用同样的推理得到同样的答案。深深吸引他的不是正确答案，而是导致这些错误答案的思维。基于这种观察结果，他提出的理论假设是，年龄大的儿童不仅在学到的东西上比年龄小的儿童多，而且他们对于问题的思考也与后者有很大的不同。这就导致他对当时流行的建立在测验分值上的智力定义产生了质疑，他认为不同年龄阶段儿童所使用的认知策略是不同的（Ginzburg & Opper，1979）。

让·皮亚杰在他后来50多年的职业生涯中，一直致力于儿童智力发展的研究。他的工作使著名的认知发展理论得以产生，该理论无可辩驳地解释了人类获得复杂思维的过程。该理论认为，儿童的发展都要经过四个认知发展阶段，其顺序相同，年龄阶段基本一致，详见表5—1。

与皮亚杰的理论本身同样重要的是他研究儿童思维能力所使用的技术。在比纳实验室里，他意识到如果要建立起新的智力概念，那么就要在研究方法上有所突破。与以往那些死板的标准化测验不同，他提出了一种新的访谈技术，即根据儿童的不同回答来改变提出的问题。通过这种方法来探测隐藏在答案背后的思维过程。

表5—1 皮亚杰的认知发展阶段

阶段	年龄范围	主要特征
感觉运动阶段	0~2岁	· 所有的知识都是通过感觉和运动获得的（例如观察和抓握） · 思维与身体运动是同步进行的 · 客体永久性概念开始形成
前运算阶段	2~7岁	· 思维与运动分离，并且思维的速度大大加快 · 能够进行象征性思维 · 非逻辑的“不可思议的（magical）”思维 · 万物有灵论：所有客体都有思想和感觉 · 自我中心式思维：不能从他人的角度看世界
具体运算阶段	7~11岁	· 逻辑思维开始发展，包括能对客体进行分类以及根据数学法则进行运算，但其对象仅限于现实具体客体 · 液体、面积和体积的守恒 · 能够推断他人的感受或想法
形式运算阶段	11岁以上	· 逻辑思维拓展到假设概念和抽象概念 · 能够借用隐喻和类比进行推理 · 能够探讨价值观、信仰以及哲学问题 · 能够思考过去和将来 · 每个人所达到的形式运算阶段的水平不同，有些人甚至根本达不到形式运算阶段

让·皮亚杰的研究最引人注目的特点之一就是他通过研究自己的孩子吕西安娜（Lucienne）、杰奎林（Jacqueline）和洛朗（Laurent）而得出了许多结论。按照今天的科学标准衡量，这种方法将会受到极大的质疑，因为它极有可能导致偏见、缺乏客观性。然而，事情总有出人意料的时候，皮亚杰通过对自己孩子的研究所得的结果已成功地应用于世界上大多数儿童。

很明显，由于篇幅有限，我们无法在这短短的一节中对他的工作细节作进一步的阐述。因此，我们将重点来讨论他对关键智力技能“客体永久性”（object permanence，也译作“客体恒常性”）概念的探索。这一能力是皮亚杰最重要的发现之一，同时对它的探讨将为我们提供更多机会来了解他的研究方法。

客体永久性的意思是即便物体不在我们的感知范围之内，我们也认为它是客观存在的。倘若现在有个人来到你面前，从你的手中把这本书取走，又跑进了另一间房间，你会认为这本书或拿书的这个人已经不存在了吗？当然不会。这本书和这个人已经在你的脑海中形成了概念，即便你已看不见、摸不到他们，你也知道他们依然存在。但皮亚杰认为，这种能力不是人生来就有的。他以实验证明，对客体永久性的认识是人在8个月大的时候才开始发展的。这种能力非常重要，因为倘若没有它，我们就无法进行问题解决和内部思维。所以，当一名儿童从感觉运动阶段（0到2岁；参见表5—1）进入前运算阶段时（2到7岁），儿童必须具备客体永久性的能力。

方法与结果

皮亚杰用非结构式评价方法研究了客体永久性这一认知技能的发展过程。由于观察对象是婴幼儿，皮亚杰的研究常以游戏的形式出现。在这些游戏中，他与他的孩子们一起玩耍，通过对他们问题解决能力以及在游戏中所犯错误的观察，皮亚杰发现在感觉运动阶段之中还有六个小阶段，这六个阶段与物体概念的形成有关。为使你更好地领略他的研究风格，下面我们将对这六个阶段作一个简要介绍，并穿插一些皮亚杰观察日记中与此相关的案例。

阶段1（出生~1个月）

在此阶段中，我们能观察到婴儿对喂养和接触的行为反射，但没有任何与客体永久性有关的迹象出现。

阶段2（1个月~4个月）

在第二阶段中，仍然没有出现与客体永久性概念有关的任何迹象，但有些行为被皮亚杰认为都是客体永久性概念的前期准备。婴儿开始有目的地重复以自己身体为中心的各种动作。例如，如果婴儿的手偶然碰到了自己的脚，他也许便会反复做出同样的动作以使这种现象反复出现，皮亚杰将其称为“初级循环反应”。在这一阶段，婴儿还可以用他们的眼睛追随物体。通常，当一个物体离开他们的视野时，他或她的视线将会继续停留在物体消失的那个点上，好像希望这个物体能再次出现。这种现象似乎是客体永久性概念的一种表现，但皮亚杰并不这样认为，因为这时的孩子还不会去主动寻找消失了的物体，如果物体不再出现，他们将会把注意力转到别的物体上，皮亚杰把这种行为称作“被动期待”。下面是皮亚杰与他的儿子洛朗之间一次互动的情况，它可以说明以上这一点。

观察2 洛朗两个月大时，我透过摇篮的顶棚观察他，我总在某个固定点出现，当我离开他的视线时，洛朗就盯着那个点看，急切地希望我再次出现（P.9）。

孩子的目光仅限于物体消失的那个地方：如果没有别的东西再次出现，那么在他的脑海里只会留下对物体知觉的一些感受，他不会再去探寻物体的去向。如果他的头脑中有物体的概念……他便会主动去各种可能的地方寻找物体的下落……但正是在这一点上他是无能为力的，因为对他而言，消失的物体还不是“永久的客体”，它仅仅是一个表象，一旦消失就无迹可寻，有时却又莫名其妙地出现（P.11）。

阶段3（4个月~10个月）

在这个阶段，孩子们开始有目的地反复操纵在环境中偶然遇到的物体（二级循环反应）。他们开始伸出手来力图抓住那些东西，用力摇它们，把它们拿到眼前仔细观察或放进嘴里。同时，孩子们的快速眼动能力也开始发展，他们的眼睛能追踪迅速移动或落下的客体。在这个阶段的后期，首次出现了“客体永久性”的信号，例如，如果孩子们看见了物体的一小部分，那么他们便会开始寻找那些在视线中还很模糊的物体。

观察23 在吕西安娜9个月大时，我给了她一只她以前从未见过的赛璐鹅。她立即抓住它，将它仔仔细细地研究了一遍。我把赛璐鹅放在她的旁边，当着她的面把它盖住，有时候盖住全部，有时候露出鹅的脑袋。吕西安娜作出了两种截然不同的反应……倘若鹅在视野中完全消失，即便吕西安娜马上就要抓住它了，她也会立即停止对鹅的搜寻……但倘若将鹅嘴露出来，她不仅会抓住看得见的部分，把动物拽到她面前，而且……有时候为了要抓住整只赛璐鹅，她会预先揭开用于遮挡的布……即使她已经在出现鹅嘴的情况下数次将布掀起，吕西安娜也从来不试着在鹅完全藏起来时，掀起那用于遮挡的布。这就证明了对整体的重新组合要比寻找看不见的东西容易得多（P.29-30）。

然而，皮亚杰仍然坚持认为，物体的概念还未完全形成。对于这个阶段的儿童而言，物体的存在并不具有独立性，它是与儿童自己的行动及感知觉联系在一起的。换言之，“儿童认为物体只露出一部分的原因是由于它们正在消失，而不是被其他物体所掩盖”（P.35）。

阶段4（10个月~12个月）

在第三阶段的最后几周与第四阶段早期，儿童已经知道即使客体不在视线之内，它们依旧存在。儿童会想方设法地主动寻找完全被隐藏的客体。从表面上看，这似乎标志着客体永久性概念已经形成，但皮亚杰认为，这种认知技能尚未得到全面发展，因为儿童仍然不具备理解“可见转移”的能力。为了便于理解，皮亚杰引用了下面的例子（你可以自己尝试一下）：你与一个11个月大的孩子坐在一起，把一个玩具完全藏在毛巾下（位置A），孩子会从毛巾下找出那个玩具。在孩子看来，客体很明显地存在着，并没有消失。然而，如果你再当着他或她的面把物体藏在毡子下（位置B），孩子将有可能返回先前发现过玩具的位置A进行搜寻。此外，你可以多次重复这个过程，孩子会重复地犯同样的错误，我们称此现象为“A非B效应”（A-not-B effect）。

观察40 杰奎林10个月大时，我让她坐在床垫上，我从她的手中取走鹦鹉，并连续两次藏在她左边的垫子下（位置A），她两次都找到鹦鹉并抓在手里。然后，我又从她手中取走鹦鹉，在她面前慢慢地把鹦鹉移到她右边的床垫下（位置B）；杰奎林非常专注地看着这个移动过程，但是当鹦鹉在B位置消失以后，她却转向左侧，到鹦鹉以前消失的那个地方（A）去寻找。

对出现在阶段4中的错误，皮亚杰作出了如下解释：这并不是由于孩子们心不在焉，而是由于他们脑中的客体概念与你我脑中的有所不同。对于10个月大的杰奎林来说，她的鹦鹉并不是一种与独立于她的行为的永恒存在物。我们先把鹦鹉藏起来，然后儿童在位置A找到了它，于是鹦鹉的概念就变成了“在A位置的鹦鹉”，这一定义不仅依赖于鹦鹉本身，而且还依赖于它所藏的地方。换句话说，在儿童的脑海中，鹦鹉仅仅是整个画面中的一部分，而不是一个单独存在的客体。

阶段5（12个月~18个月）

大约从1岁左右开始，儿童获得了追踪物体连续可见转移的能力，并且能够在物体最后出现的地方找到它。出现这种现象后，皮亚杰认为，孩子便进入了感觉运动阶段的第五阶段。

观察54 我们让11个月大的洛朗坐在A和B两个垫子中间。我在A与B两个位置之间交替着隐藏手表；洛朗不断地在手表最后出现的地方进行搜寻，有时候在位置A，有时候在位置B，而不像在前一个阶段那样，总是在第一次手表消失的位置寻找（P.67）。

然而，皮亚杰指出，真正的客体永久性概念仍未完全形成，因为儿童还不能够理解被皮亚杰称为“不可见的转移”的现象。设想一下下面的例子：你看见一个人把一枚硬币放在一个小盒子里，然后他背对着你走到梳妆台前，打开了抽屉；当他回来的时候后，你发现那个盒子里空空如也，这就是所谓的“不可见的转移”。当然，作为成人的你会自然而然地走到梳妆台前查看抽屉。但正如皮亚杰所证明的，这种能力也许亦不是天生的。

观察55 18个月大的杰奎林坐在一块绿色的小毯子上，高高兴兴地玩弄一个土豆（对她来说，土豆是一个新玩意儿）。她把土豆放在一个空盒子里，又把它拿出来，玩得不亦乐乎。然后我当着她的面把土豆拿过来，放进盒子里，然后我把盒子放在毯子下面，并把土豆倒出来，把它藏在毯子下，最后取出空盒子，我没有让杰奎林看见我玩的小伎俩。虽然杰奎林一直盯着毯子，也知道我在毯子下面做了点手脚，可当我对她说："给爸爸土豆"时，她开始在盒子里寻找土豆，还抬头看看我，又看了一会儿盒子，再看看毯子……但是，她并没有掀起毯子去寻找下面的土豆。在此后连续五次的实验中，得出的结果都是这样（P.68）。

阶段6（18个月—24个月）

最后，孩子们将进入感觉运动阶段的末期（见表5—1），这时客体永久性概念就彻底形成了。进入这个阶段的标志是他们能找出经过"不可见的转移"的东西。

观察66 杰奎林1岁零7个月时，已有能构想物体被隐藏在重重屏蔽物之下的能力……我把铅笔放在盒子里，用一张纸将盒子包起来，再用手帕扎裹一层，最后用贝雷帽和床单把它罩起来。杰奎林先揭开贝雷帽和床单，然后再解开手帕，却没有立即发现盒子，但是她继续寻找，显然她已确信盒子的存在。然后她觉察到了纸，并立即明白了其中的奥妙，她撕开纸，打开盒子，找到了铅笔（P.81）。

皮亚杰认为，客体永久性这种认知技能是真正思维的开始，是运用洞察力和符号来解决问题的能力的开始。这就为儿童进入下一个阶段（前运算阶段）的认知发展做好了准备。在前运算阶段，思想与行动相对独立，使思维的速度能显著提高，换句话说，客体永久性概念是所有智能的基础。正如皮亚杰所说：

在众多的事物当中，客体守恒是客体定位的机能。也就是说，儿童既能明白当物体消失时，它依然存在；也能理解客体去往何处。这一事实表明，客体永久性的图式建构是同现实世界的整个时空组织和因果关系密切联系在一起的（Piaget & Inhelder，1969）。

讨论

这种操作和行为观察的方法构成了皮亚杰工作的基础，并贯穿其对认知发展四阶段的建构。皮亚杰坚信他的阶段理论可广泛地适用于世界上所有的儿童，无关乎文化和家庭背景的差异。另外，他强调了在感知运动阶段中与客体概念发展有关的几个重要方面（参见Ginzburg & Opper对此问题的详细叙述，1979）。

1.每个阶段的年龄范围仅是一个近似值。因为皮亚杰的早期工作只涉及三个孩子，要很确定地预测年龄范围，这对皮亚杰而言相当困难。例如，他在杰奎林1岁零7个月时观察到的能力，吕西安娜在1岁零3个月时就已表现出来了。

2.皮亚杰坚信各阶段的发展顺序是固定不变的。所有儿童在进入下一个阶段时，必须经历此前的每一个阶段，不可能有任何跳跃发生。

3.从一个阶段到另一阶段的变化是随着时间推移而逐渐发生的。这样，在前一阶段的错误就会随着新心智能力的成熟而渐渐减少。皮亚杰认为，孩子们有时会处在两个阶段之间，有时也会同时表现出滞后或超越某个阶段的一些能力，这都很正常、很普遍。

4.儿童的认知发展进入下一个更高级阶段后，与前一阶段有关的行为并没有完全消失。对于一个处于第六阶段的儿童来说，采取第五阶段的智力策略来解决问题是很自然的。若这些策略未获成功，儿童便会采用新的方法，即用第六阶段典型的推理方法来做新的尝试。

批评与近期应用

尽管皮亚杰的认知发展概念在发展心理学领域占统治地位已将近40年，但其间也确实遭到一些批评。有人把质疑的目光集中在皮亚杰关于认知发展阶段的划分上。许多学习理论学家不同意皮亚杰在这个问题上的观点，皮亚杰认为认知的发展过程由不连续的阶段组成，而他们则坚决主张智力的发展是一个连续的过程，在这个过程中没有任何特殊的顺序。他们坚信，认知能力就像其他行为一样，是模仿、学习和条件反射的结果。

另一些批评者认为，皮亚杰断言的特定能力将会出现的年龄段并不正确。甚至有些人认为，某些特定的能力也许在出生时就已出现，客体永久性就是被广泛讨论的一种。在一系列设计精巧的研究中，芮妮·贝拉吉恩（Renee Baillargeon）和她的助手采用了一种新的研究方法——注视偏好（研究5中具体论述了这种研究方法的提出），证明了在两个半月大的婴儿头脑中已经出现了客体永久性概念的最初形式（Aguilar & Baillargeon，1999；Baillargeon，1987）。她和其他学者声称，由于需要运动技能的配合，皮亚杰所用的方法不足以对幼小婴儿的能力进行准确的测量。皮亚杰的理论概念和研究发现不断地影响着其他更广阔的研究领域。一个明显的证据是，现今每年仍有超过50篇的文献引用皮亚杰著作中的内容来作为讨论的基础。例如，有一项研究用类似皮亚杰所做的游戏，比较了6个月大和半个月大的婴儿对黑暗处的物体和光亮处藏在布下的物体的搜索倾向（Shinskey & Munakata，2003）。有趣的是，研究者发现“婴儿在黑暗中寻找物体的情况下表现更好”。这种现象的原因是什么呢？其中一个解释是也许遮盖布的出现干扰了婴儿本身就很弱的在脑中表征物体的能力。另一个解释是从进化和适者生存的观点来看，比起寻找藏在光明处的物体来说，搜寻黑暗中（危险环境中）的物体的能力要更为重要。

近期有一项与皮亚杰的工作相关的研究引起了人们的注意，该研究发现婴儿区分不同物体的能力与他们理解这些物体名称的能力之间的联系（Rivera & Zawaydeh，2007）。这项研究使用注视偏好技术，发现10个月或11个月大的婴儿能够区分他们知道名称的物体。作者提出“这些结果表明，对于完形的/非完形的（隐藏的和显露的）物体名称的理解能够提供一种类似于“胶水”的东西，这种东西能够允许婴儿将物体的心理指标与其感知特征联系在一起（因而加速形成两个心理指标，而不是一个）”（P.146）。这说明，相对于物体，知道物体的名字似乎有利于婴儿存储更加独特和可识别的物体表象。

还有一项研究发现，婴儿的睡眠与其客体概念的发展有关（Scher，Amir & Tirosh，2000）。这种关系是，在同年龄组（9个月）的婴儿中，客体永久性概念水平较高的婴儿比水平较低的婴儿睡眠状况要好。如果你不能确信，第二天早晨醒来你的财物依旧原封不动（没有消失），你可能也睡不安稳吧。

结论

随着对婴儿认知能力研究的方法不断推陈出新（例如注视偏好、习惯化—去习惯化技术的应用），皮亚杰的一些研究结果也受到了质疑（关于此类研究的更多信息，请参考本书的研究5，即范茨进行的研究；也见Craig & Dunn，2007）。事实上，围绕着他的认知发展理论，现在仍存在大量争论。这种争论是非常有益的，它所引发的研究将最终推动我们对认知能力的深入了解。

尽管仍有争论出现，但对所有相关的研究领域而言，皮亚杰的理论仍然处于基础地位，催生着各种新型研究。他的工作还在继续指导、启发着人们对儿童研究、教育方法及家教风格等方面的思考。皮亚杰的贡献无论在过去，还是现在都是不可估量的。

Aguilar, A., & Baillargeon, R.(1999).2.5-month-old infants′ reasoning about when objects should and should not be occluded.Cognitive Psychology, 39(2)，116-157.

Baillargeon, R.(1987).Object permanence in 3-and-a-half-and 4-and-a-half-month-old infants.Developmental Psychology, 23, 655-664.

Craig, G., & Dunn, W.(2007).Understanding human development.Upper Saddle River, NJ; Pearson Prentice Hall.

Ginzburg, H., & Opper, S.(1979).Piaget’s theory of intellectual development.Englewood Cliffs, NJ: Prentice-Hall.

Piaget, J., & Inhelder, B.(1969).The psychology of the child.New

York: Basic Books.

Rivera, S., & Zawaydeh, A.N.(2007).Word comprehension facilitates object individuation in 10 and 11-month-old infants.Brain Research , 1146 , 146-157.

Scher, A., Amir, T., & Tirosh, E.(2000).Object concept and sleep regulation.Perceptual and Motor Skills, 91(2), 402-404.

Shinskey, J., & Munakata, Y.(2003).Are infants in the dark about hidden objects? Developmental Science, 6, 273-282.

研究19 你的品德如何

Kohlberg , L.(1963).The development of children’s orientations toward a moral order: Sequence in the development of moral thought.Vita Humana, 6 ，11-33.

你是否思考过自己的道德水平与别人相比是怎样的？在生活中你对事物作出判断的道德准则到底是什么？如果你仔细思考一下这个问题，经验将告诉你，人们在思想和行为道德准则上存在很大的个体差异。道德准则一般被心理学家定义为：帮助人们判断对和错的态度和信念。个人的道德准则是由其文化背景的行为准则和规范所决定的，而这些准则、规范已为个体内化为自己的一部分。道德准则不是与生俱来的，也就是说，你不是刚出生就有道德准则。随着从童年走向青年，再到成人，你判断对错的标准也在不断发展。但是，这种道德观念从何而来？这种将一整套文化准则内化为个体自我意识一部分的过程又是怎样的呢？

在道德形成的研究历史上，两位最著名和最有影响力的人物可能就是让·皮亚杰（参考本书第18项研究）和劳伦斯·柯尔伯格（Lawrence Kohlberg，1927年~1987年）。柯尔伯格在芝加哥大学所做的研究综合并拓展了皮亚杰关于智力发展的很多观念，且重新激起了人们对该领域的研究兴趣。柯尔伯格关心的问题是：没有是非观念的婴儿是如何形成道德准则的？

基于皮亚杰的研究，柯尔伯格提出了自己的理论主张：人类所拥有的独一无二的道德判断能力是以一种可预测的方式在整个儿童时期形成起来的。而且，他相信，与皮亚杰的智力发展阶段相似，道德准则也存在某种可区分的、特定的发展阶段。正如柯尔伯格所解释的，“只有当儿童逐渐把这些价值观与他已理解的社会秩序以及他作为一个社会自我的目标联系起来时，他才能将父母及其文化背景的道德价值观内化为他自己的一部分”（Kohlberg，1964）。换言之，儿童必须达到某种智力发展阶段才有可能达到相应的道德发展水平。

有了这些想法，柯尔伯格便着手开发一种研究儿童道德判断能力的方法。从这些研究中，他创建了他那得到广泛认可的道德发展理论。

理论假设

柯尔伯格指出道德形成遵循一定的发展阶段，他在使用“阶段”这一概念时是非常严谨和准确的。我们很容易想到所有的能力发展都有不同

阶段，但是，心理学家还是对那些随时间逐渐发生的变化（比如一个人的身高）与在不同阶段里差别很大的变化作了区分。因此，当柯尔伯格谈及“童年和青少年道德结构的发展阶段”时，他是指：（1）每个阶段都有一种独一无二的道德思维方式，而不是对成人道德概念理解的逐渐深化；（2）各阶段总是以固定的顺序出现，不可能跳过任何一个阶段，也绝对没有倒退的情况出现；（3）阶段具有优势性，即儿童理解所有处于他们现有道德阶段以下的道德判断，且至多只能对高于他们所处阶段以上一个阶段的道德问题有某种程度的理解。鼓励、教育和练习都不能使儿童向高于他们应有阶段的道德阶段发展。而且，儿童倾向于使用他们所达到的最高道德发展水平来对事物进行判断。隐含在这一道德发展阶段中的规则是：无论个体之间是否存在经验和文化上的差异，发展阶段是具有普遍性的，且它们以固定的顺序向前发展。

柯尔伯格相信，可以通过给不同年龄的儿童提供道德判断的机会来对其道德形成阶段理论加以探究。如果发现儿童做出道德决策的思维方式随年龄而有规律地发展，那么这将可以证明道德阶段论是基本正确的。

方法

柯尔伯格的研究方法相当简单。他向不同年龄的儿童提供10个假定的道德两难故事，每位儿童需接受两个小时的关于这些故事的访谈。研究者对访谈进行录音，以便对儿童所使用的道德推理方式进行进一步的分析。下面是柯尔伯格的两难故事中被引用次数最多的两则：

弟弟的难题。乔的爸爸许诺说，如果乔挣够了50美元，便可以拿这笔钱去野营。但后来他又改变了主意，让乔把所挣得的50美元都交给他。乔撒谎说只挣到了10美元，他把10美元交给了爸爸，拿另外的40美元去野营。临走之前，乔把挣钱和向爸爸撒谎的事告诉了他的弟弟阿里克斯。阿里克斯应该把事情的真相告诉他的爸爸吗（P.12）？

海因茨的难题。在欧洲，一位妇女因患有一种特殊的肿瘤而面临死亡。医生们认为只有一种药或许能挽救她的生命。那是她所在的镇上的一位药剂师最近研制的一种含镭的药物。这种药的成本昂贵，而且这位药剂师向购买者索要高出成本10倍的高价。他花了200美元制造“镭”，但在售出时，一小瓶就卖2000美元。这位病人的丈夫叫海因茨，他向他认识的所有人都借了钱，但在最后他也只能借到1000美元，仅仅够要价的一半。他向药剂师恳求说他的妻子快死了，求求他便宜一点卖给他或者允许他以后再支付另一半的钱。但药剂师却说：“不行，我研制该药的目的就是为了赚钱。”因此，海因茨绝望了，他闯进了药店，为他的妻子偷了治病的药。海因茨应该这样做吗（P.17）？

柯尔伯格最初的被试是居住在芝加哥郊区的72名男孩。这些男孩分属于3个年龄组，即10岁、13岁和16岁。每个年龄组中有一半被试来自社会经济条件中下的家庭，而另一半则来自社会经济条件中上的家庭。在两个小时的访谈中，这些孩子表达的道德观点从50到150个不等。

下面是柯尔伯格所引用的四名不同年龄的儿童面对道德两难处境时所作的反应：

丹尼（Danny），10岁，就“弟弟的难题”作出以下回答：“一方面，他应该告诉爸爸事情的真相，否则的话，他的爸爸或许会生他的气，甚至会打他的屁股。另一方面，也许他应保持沉默，否则他的哥哥会揍他。”（P.12）

唐（Don），13岁，就“海因茨难题”作出以下回答：“药剂师的错。他是不公道的，索要高价且不顾别人的死活。海因茨爱他的妻子，并想救她。我认为任何人都会这么做的。我相信他不会被关进监狱。法官会全面看待这场官司，并明白药剂师是在漫天要价。”（P.19）

安迪（Andy），13岁，就“弟弟的难题”作出以下回答：“如果我爸爸事后发现了真相，他将不会再信任我；（反之）我的哥哥将不再信任我，但如果我不说出真相，我不会觉得太糟。”（P.20）

乔治（George），16岁，就“海因茨的难题”作出以下回答：“我不这样看，因为药剂师有权决定药的价格。我不能说海因茨确实是对的，尽管我猜想任何人都会为了妻子而这么做。他宁愿进监狱，也不愿看到他的妻子死去。在我看来，他有正当的理由这么做，但从法律的角度看，他是错的。至于究竟是对是错，我不能发表更多的意见。”（P.21）

基于这些陈述，柯尔伯格和他的同事界定了六个道德发展阶段并把孩子们的所有陈述分别归入其中某一个发展阶段。此外，被试作出道德判断的动机也有六种类型，它们与每个阶段一一对应。值得注意的是，柯尔伯格所描述的每个道德推理阶段都可以普遍适用于儿童可能面临的任何情境。尽管道德发展阶段无法预测一个儿童面对真实的两难处境时所采取的特定行动，但却能预测该儿童在决定一个行动所进行的推理过程。

结果

柯尔伯格把这六个发展阶段分成三种道德水平，具体参见表5—2。道德观念发展的早期被柯尔伯格称为“前道德水平”，该水平的特征是以自我为中心，看重个人利益。它包括最初的两个阶段：在第一阶段，儿童认识不到他人的利益，其道德行为是出于对不良行为将要受到惩罚的恐惧；在第二阶段，儿童开始意识到别人的利益和需要，但他们的道德行为是为了别人回报以同样的道德行为。这时，良好行为的本质是儿童为了满足自身需要而对情境施行的控制。

表5—2 柯尔伯格的道德发展六阶段

水平 1 前道德水平
阶段 1. 惩罚和服从定向（行为的后果作为是非标准）
阶段 2. 朴素的利己主义定向（以个人需求的满足与否决定事情的好坏）
水平 2 习俗角色遵从的道德
阶段 3. 好孩子定向（取悦于别人就是好的）
阶段 4. 维护权威定向（维护现有的法律和社会秩序，尽职尽责就是好的）
水平 3 自我接受的道德准则的道德水平
阶段 5. 墨守成规和契约定向（社会价值和个人权利作为是非标准）
阶段 6. 个人的良心和原则定向（是非判断基于依照普遍原则所确立的个人哲学）

（引自 P.13）

在道德发展到第二种水平时，习俗道德开始作为人际关系中个体角色认知的一部分发挥作用。在第三阶段，儿童的道德行为是为了达到他人对自己的期望并维持与他人之间的信任和忠诚的关系。按照柯尔伯格的观点，正是在此期间，黄金规则思维（golden rule thinking）开始出现，儿童开始关注其他人的感受；到了第四阶段，儿童开始意识到法律和秩序的存在，并表现出对法律和秩序的尊重。在此阶段的儿童从更大的社会系统的角度看待事物，并以行为是否遵纪守法为衡量好公民的尺度。他们对现有的社会秩序表示认同，并认为遵守法律的行为都是好的。

当一个人进入第三种水平时，他的道德判断开始超越现有的法律。在第五阶段，人们开始承认某些法律比另外一些法律好。现实中有合情不合法的事，也有合法不合情的事。处于此阶段的个体仍坚信，为了维护社会和谐，人们应该遵守律法，但他们也会通过特定的程序寻求对法律的修正。柯尔伯格声称，这一阶段的人在尝试整合道德和法律时将面临冲突。

最后，如果一个人达到了第六阶段，他或她的道德判断将建立在对普遍道德行为准则的信仰之上，当法律与道德准则相冲突时，个体将依据他/她的道德准则做出决策而不考虑法律。决定道德的将是个体内在的良心。柯尔伯格在本篇论文的研究以及以后的其他研究中均发现，只有极少数人能够完全达到第六阶段。他最后认为这种道德推理水平只能在具有道义感的伟大领导者身上发现，比如甘地、梭罗以及马丁·路德·金等。柯尔伯格认为：

> 道德的动机层面可以用被试在论证道德行为时所提到的动机加以界定。六种动机水平是彼此独立的，每一种动机都对应一种道德发展阶段，具体如下：（1）被他人惩罚；（2）对物品的占有或得到他人的奖励；（3）受到他人的谴责；（4）被公认的权威斥责并伴有罪恶感；（5）所在群体的尊重和不敬；（6）自我谴责。（P.13）

最后，柯尔伯格的道德阶段论至关重要的一点是：各种水平的道德推理是随着年龄的增长而发展的。为了验证这一观点，柯尔伯格将不同年龄儿童对两难问题的回答，按道德水平的发展阶段进行了分类。图5—2概括了这些发现。很明显，随着年龄的增长，儿童更多地使用高级阶段的道德推理来回答问题。其他的统计分析表明，学会使用每一阶段

的道德推理能力对于更高一级道德阶段的发展而言是必不可少的。

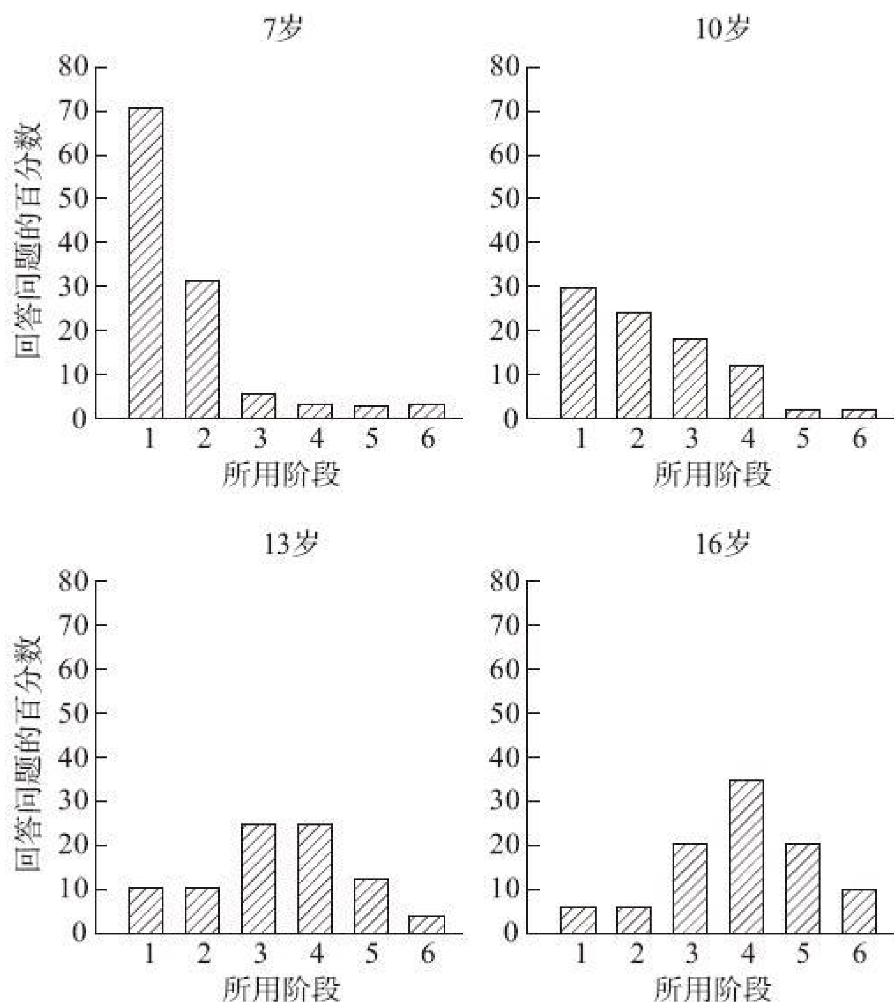


图5—2 道德推理发展的年龄阶段

注：柯尔伯格解释，打*号的7岁组的数据来自12名附加被试。

资料来源：引自实验研究论文的第15页。

讨论

在柯尔伯格就他的研究结果所做的讨论中，他指出，这一新的概念体系阐明了儿童是如何主动地以一系列可预测的有序的阶段来对他周围的世界进行道德构建的。对儿童而言，这一过程不应被简单地视为成人通过口头解释和惩罚，将成人的道德准则同化和内化为儿童的一部分，

而是一种儿童在与社会和文化环境的相互作用中发展起来的道德认知结构。按照这一观点，儿童不是在学习道德标准，而是在建构道德标准。这就意味着一个儿童在完成第一阶段和第二阶段的道德认知建构前，是根本无法理解和使用第三阶段的道德推理的。同理，一个人除非已经经历而且建构了前四个阶段的内在道德模式，否则他是不会使用属于第五阶段的基本人权的概念来解决道德两难问题的。稍后，我们将对柯尔伯格在该研究及其后续研究中所涉及的更深含义进行讨论。

批评意见及近期应用

自这一开创性研究以来的30年间，柯尔伯格对道德发展阶段论不断地拓展和修正的，同时也遭到了来自各方面的批评。其中最常被引用的一种批评是，即便柯尔伯格关于道德判断的观点是正确的，也并不意味着这些判断能被运用于道德行为。换言之，一个人心里想的和口头上说的或许不会反映在一个人的真实行动之中。很多研究已经证明，在道德判断和道德行为之间缺乏一致性。然而，另一些研究却发现二者之间确实存在关联。与此有关的一种有趣的研究方向认为，柯尔伯格在确定某人是否将依据他或她的道德判断而行动时，情境因素的重要性被大大低估了（参见Kurtines，1986）。该批评显然有一定的道理，柯尔伯格也承认他的理论只适用于道德判断。但柯尔伯格指出，情境力量或许有时会改变道德行为的事实，但这并不能否认柯尔伯格所描述的道德判断的发展进程。

另一种对柯尔伯格的工作的批评针对的是柯尔伯格道德推理六阶段的普遍适用性。这些批评者声称，柯尔伯格的六阶段论所提出的道德解释，只符合西方的个人主义社会。因此，它不适用于占世界人口大多数的非西方文化（参考本书研究28，Triandis关于不同文化之间的讨论）。为了捍卫柯尔伯格的道德发展阶段具有普遍性的观点，塞纳瑞（Snarey，1987）在27种不同的文化背景中完成了45项研究。在每项研究中，研究者都发现了他们的被试以相同顺序通过各阶段，没有颠倒现象的出现，而且第一阶段到第五阶段在被研究的所有文化中都普遍存在。不过有趣的是，在一些文化中（如中国台湾、巴布亚—新几内亚以及以色列等），某些道德判断无法纳入到柯尔伯格的六阶段中的任何一个中去，这些道德判断是建立在全社会的利益基础之上的。这种推理在美国被试中是无法找到的（参考本书研究28，Triandis关于个人主义与集体主义文化的研究）。

最后，第三方面的批评是基于这样的想法：柯尔伯格的道德发展阶段论或许不能同等地适用于男性和女性。对此提出质疑的代表人物是卡罗尔·吉利根（Carol Gilligan，1982）。她认为，女性和男性思考道德问题时所用的方式并不完全相同。她在自己的研究中发现，与男性相比，在做出道德决策的过程中，女性更多地谈论人际关系、对他人的责任、避免伤害别人以及人们之间保持联系的重要性。她把女性道德判断

所依赖的基础叫作关怀定向（care orientation）。基于这种性别差异，吉利根认为，女性之所以在柯尔伯格的测验中得分偏低，是由于较低的道德发展阶段中涉及更多人际关系问题（如第三阶段，它主要建立在形成人与人之间信任和忠诚的关系的基础上）。另一方面，吉利根认为，男性是基于正义而作出判断的，而这更符合柯尔伯格所提出的高级发展阶段。她主张，这些道德取向并没有优劣之分，如果女性被判定为比男性处于更低的道德水平，那只是因为柯尔伯格的道德阶段论存在一种无意的性别偏见。

研究者大多无法为吉利根的观点提供证据。很多研究发现，使用柯尔伯格的研究方法并没有发现在道德推理上存在性别差异。对此，吉利根承认，尽管女性能够使用所有的水平进行道德推理，但在她们的现实生活中，她们不会作出相应的行为。相反，正如前面所讨论的，女性看重的是人际关系层面，这已经被研究所证实。一项研究表明，女孩愿意付出更多的努力去帮助另一个需要帮助的人，她们在关于移情的测验上倾向于获得更高的分数（参见Hoffman，1977年针对性别话题所作的全面讨论）。

柯尔伯格早期对道德判断的发展所做的研究不断被许多学科领域的研究所引用。不出所料的是，目前在法律和犯罪审判领域有相当一部分研究使用了柯尔伯格的理论模型。一项基于柯尔伯格理论的研究调查了妇女在怀孕期间酗酒对其孩子道德发展的影响（Schonfeld，Mattson & Riley，2005）。怀孕期间酗酒抑制儿童智力发育已有十分明确的证据，这项研究也发现那些在出生前接触酒精的儿童和青少年的道德成熟度低于对比组。根据柯尔伯格关于道德发展的阶段理论，当那些“接触酒精”个体正处于关注将自己的损失最小化时的初级阶段（如第二阶段），而对照组的个体已经发展到了关注他人和社会规范的阶段（第三阶段）（P.550-551）。

另一项引用柯尔伯格理论进行的研究考察了儿童所提供的目击者证词的准确性（Bottoms等，2002）。一些3~6岁的儿童在母亲的陪同下参加一个游戏活动。一半儿童被告知不可以玩室内的某些玩具。然而，当研究者离开后，这些儿童的母亲就鼓动他们去玩那些“禁止”的玩具，但是告诉他们“要保密”。一段时间后，研究者开始询问他们是否玩过被禁止的玩具。结果表明，“对于年龄较大的儿童来说，被告知保守秘密的儿童比没被告知保守秘密的儿童隐藏了更多信息。而对于年龄较小的儿童来说，保密的要求对他们陈述的内容不会有什么显著的影响”（P.285）。很多时候，儿童被要求为成人违法或有害的行为保密。了解多大的儿童能够领会保密的含义并能保守秘密，对于在诉讼中儿童目击证词的使用可能起到重要作用（见本书中Loftus关于目击证词的研究）。

结论

在行为科学领域内针对柯尔伯格的研究工作所展开的对话和争论持续至今（例如，Goodwin & Garley，2010），且明显呈现出继续下去的迹象。该理论的最终价值和重要性有待于人们做出更清晰的界定。但不管怎样，极少有关于人类发展的概念能像柯尔伯格的道德发展阶段论那样，引发如此众多的研究、思考和争论。而该理论的社会价值似乎已在柯尔伯格1964年所做的一段陈述中有所显现：

尽管任何道德教育的概念都认识到，父母不可避免地会把自己的行为准则和道德判断强加于儿童，但把道德教育主要限定在激发儿童自身道德判断的产生和对行为的控制却是可能做到的。我已经发现，有些教师劝诫13岁的少年不要作弊时是这样说的："因为你抄袭的那个人的回答或许是错的，并且这样做的话对你没有任何好处。"这些少年中的大多数人后来都能提出更好的不作弊的理由……儿童很可能会抛弃那些低于他们道德发展水平的道德推理，他们同样也无法同化那些远远超出其道德发展水平的道德推理（P.425）。

Bottoms, B., Goodman, G., Schwartz-Kenney, B., & Thomas, S.(2002).Children's use of secrecy in the context of eyewitness reports.Law and Human Behavior, 26, 285-313.

Gilligan, C.(1982).In a different voice: Psychological theory and women'development.Cambridge, MA: Harvard University Press.

Goodwin, G.P.& Darley, J.M.(2010).The perceived objectivity of ethical beliefs: Psychological findings and implications for public policy.Review of Philosophical Psychology，1，161-188

Hoffman, M.L.(1977).Sex differences in empathy and related behavior.Psychological Bulletin, 84, 712-722.

Kohlberg.L.(1964).Development of moral character and moral ideology.In H.Hoffman & L.Hoffman (Eds.), Review of child development research (Vol.1).New York: Russell–Sage Fanndation.

Kurtines, W.(1986).Moral behavior as rule-governed behavior: Person and situation effect on moral decision making.Journal of Personality and Social Psychology, 50, 784-791.

Schonfeld, a., Mattson, S., & Riley, E.(2005).Moral maturity and delinquency after prenatal alcohol exposure.Journal of Studies on Alcohol 66(4), 545-554.

Snarey, J.(1987).A question of morality.Psychological Bulletin, 97, 202-232.

研究20 让你愉快的控制力

Langer, E.J., & Rodin, J.(1976).The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting.Journal of Personality and Social Psychology, 34，191-198.

“控制力”这个概念在心理学中看似微不足道，但它却是一个能对人类所有的行为产生极为重要影响的因素。我们在这儿所谈论的控制力并不是指控制别人的那种能力，而是指你所拥有的控制你自己生活和事件的个人力量。这种能力与你对自我力量的感觉以及在各种情形下个人选择的有效性有关。人们大多认为自己至少能控制其个人的意愿。生活中，你做过不少选择，有好有坏，正是它们引领你成为今天的你。也许你未曾认真地思考过这些，事实上你在生活中所做的选择比你想到的要多得多，你每天都在选择或决定自己的行为。当你的控制力受到威胁的时候，你就会出现负面情绪（生气、狂怒、愤慨），并会以某种行为来反抗，以此来恢复你的个性自由。如果有人告诉你，你必须做某件事，你很有可能就会拒绝他，或是朝相反的方向去做，这已是太平常的事了。或者，反过来讲，如果我们禁止某人做某些事，他们反倒会觉得被禁止的活动比以前更有吸引力（还记得罗密欧和朱丽叶吗），这种反对任何限制我们的自由、阻止我们的愿望的倾向被称为“阻抗”（reactance）。

如果我们对控制个人环境的需要与人类本性同样重要，倘若你丧失了控制力，且再也无法将它找回来，你认为结果会怎样？你很可能会经历以焦虑、气愤、狂怒、沮丧、无助甚至是以生理疾病为形式的压抑状态。研究表明，当人们处在压力情境中时，如果他们相信自己对这一应激事件能够有所控制，那么压力情境的负面影响就会减少。例如，在拥挤的电梯里，有人愿意站在靠近控制板的地方，那样他们会觉得电梯不那么挤，也不感到焦虑。不管他们是否掌控“逃跑”，他们认为自己比这个环境中的其他人更具有控制力（Rodin，Solomon，& Metcalf，1979）。另外一项众所周知的研究表明，当人们发现他有能力控制住一种紧张情况时，那他的压力会减少很多（参考Glass & Singer，1972）。例如，让被试暴露于突然的强噪音中，然后要他们完成问题解决的任务。其中一组被试不能控制噪音，而主试告诉另一组被试，他们可以在任何时候按压按扭以停止噪音。然后，又嘱咐他们如果可以不按压的话就尽量别按压。没有控制力的那组被试任务完成情况明显要比那些相信自己可以对噪音进行控制的被试差得多。另外，在后一组中实际上没有被试按压过按扭，所以他们暴露于噪音的时间与无控制力的那组被试一样多。

所有这些归结起来就是，当我们拥有控制力时，我们便是更快乐和

更有成就的人。不幸的是，在我们的社会中，许多人都会经历这样一个阶段，他们丧失了这种控制力量，连对自己的生活做出最简单的选择也会受到限制，这个生活阶段就叫作“老年”。我们许多人都曾听说过或直接接触过健康状况突然恶化的老年人，他们对退休和疗养院都十分警惕，因为这时他已经被当作一个老年人了。许多疾病如结肠炎、心脏病以及抑郁症等都与在疾病发生之前所产生的无助感、失控感密切相关。老年人进入疗养院后，他们必须忍受的最困难的转变之一就是，放弃对自己日常生活的控制力，这便影响了个人对自己命运的选择。在我们思考这个研究之前，兰格和罗丁一直在研究控制力的问题，他们决定把这些观点放进真正的养老院环境里来进行检验。

理论假设

如果一个人在生活中丧失了对自己负责的能力，往往会引起他的不快，并危害他的身体状况，那么提高人的控制力量就会有截然相反的结果。兰格和罗丁想通过增进一组疗养院老人对自己负责的能力和选择权，来直接验证上述理论观点。以前人的文献和自己的前期研究为基础，他们预测如果给病人以控制感，将明显提高他们的心理警惕性、活动水平、生活满意度以及其他可测量的行为和态度。

方法

被试

兰格和罗丁与康涅狄格州的一个名叫阿登屋（Ardn House）的养老院进行合作。这个疗养院被康涅狄格州评为这个地区最好的疗养单位之一，它能提供高质量的医疗服务、娱乐设施和居住条件。该养老院共四层，规模大而又现代化。总体而言，这个疗养院里的老人身心健康状况相近，有相似的社会经济背景。新成员进入疗养院后，就被安排在一间基本设施齐全的房间里。因而，整个疗养院里，老人的基本状况大体而言是相同的。随机挑选两层楼，让住在这两层楼里的老人分别接受两种实验处理。四楼的老人（8男，39女）接受了“责任感提升”的训练，二楼的老人则作为对照组（9男，35女）。91名被试的年龄跨度为65岁到90岁。

过程

疗养院的管理人员同意辅助研究人员在两种实验条件下进行工作。管理人员是一名33岁的男性，他性格外向、友好，每天都要与老人们打交道。他给那两层楼的老人开了个会，告诉他们一些有关疗养院的新信

息。他把两条信息分别传递给了这些老人，表示疗养院会把他们的生活安排得尽可能舒适而令人满意，还说明了他们能够享受的几种服务。但两组被试得到的信息是有重要差别的。

“责任感提升”组（4楼的老人）得到的信息是，他们有照顾自己的责任，并有权决定如何安排自己的时间。他的具体说明是这样的：

> 你们可以决定自己房间的设施布置——无论你是希望它就像现在这样，还是希望工作人员帮你再重新布置一下……你有责任让我们知道你的意见，告诉我们你想做什么样的改变，告诉我们你所希望的事情。另外，我想利用这个机会送给你们每人一个由阿登屋疗养院准备的礼物（工作人员拿着装满小植物的盒子来到老人面前。所有老人要做两个决定：第一，他们是否想要植物；第二，选择一种自己喜欢的植物。结果，所有的老人都给自己选了一种植物）。这些植物是你们的了，请你们自己照顾好自己的植物。
>
> 最后还有一件事我想通知你们，那就是下周四、周五的晚上我们将各放映一场电影。如果你想看的话，请在两天之中选择一天（P.194）。

对控制组（2楼的老人）说，疗养院多么希望使他们的生活更充实、更有趣。他的说明如下：

> 我们希望你们的房间能尽可能地舒适，并且我们也已尽力为你们作了这样的安排。我们希望你们在这儿能感到高兴，我们的责任就是给你创造一个幸福的家，让你能为它而感到自豪，我们将尽全力在各个方面帮助你们。
>
> 另外，我想利用这个机会送给你们每人一个由阿登屋疗养院准备的礼物（护士拿着装小植物的盒子走了一圈，发给每位老人一棵植物）。这些植物就是你们的了，护士会每天替你们给植物浇水照顾它们。
>
> 最后还有一件事我要通知你们，那就是下周四、周五的晚上我们将各放映一场电影。稍后将会通知你们安排哪一天去看（P.194）。

三天之后，这名管理人员到每位老人的房间里又去了一次，并重复了同样的信息。

不难看出，这两种信息之间存在着重要的差异：实验组的被试（四楼的老人）在生活中有进行选择的机会，用多种方法调节自己的生活；而对于控制组的被试（二楼的老人）而言，虽然其他因素与实验组基本相同，但他们收到的信息是，将由疗养院来代他们做出大部分决策。这一实验过程持续了三个星期（应该注意的是，四楼老人所拥有的控制力水平，是所有居住在这个疗养院中的人都能获得的。在这个实验中，研究者只是对实验组重申他们所拥有的这一控制力，并使他们更清楚这一点）。

测验结果

在这项研究中，研究者运用了数种测量方法（因变量）来了解不同的责任状态是否会产生不同的结果。他们运用两份问卷进行了两次施

测，一次是在管理员召开会议前的一周，另一次是在召开会议的三周后。其中一份问卷由老人们填写，主要内容涉及对自己控制力的评估，及在疗养院生活的快乐程度和活动水平。另一份问卷由各层楼的护士（她们并不清楚研究者在进行实验）填写，问卷要求她们用10点量表评估老人们在快乐、机敏、依赖、社交、活力等方面的水平，并记录下他们的饮食和睡眠习惯。另外，还有两个测验测量的是老人的实际行为表现：其中一项记录的是老人在第二周出席电影观摩的情况；另一项以竞赛活动的形式出现，研究者让老人们猜测在一个大坛子里共有多少粒果冻豆，想参加的人只需在一张纸条上写下他们的名字和猜测结果，并将纸条放入坛子旁边的一个盒子里即可。

结果

表5—3展示了两份问卷的调查结果。我们可以很清楚地看出，两组被试的差异是非常明显的，其结果支持了兰格和罗丁的预测，即选择的机会和个人的控制力能产生积极的作用。责任感提升的那组被试报告说他们觉得自己比对照组的被试更快乐，更富有活力，而且访谈者对他们机敏程度的评定也高于对照组的评定，所有这些差异在统计上都具有显著性。更显著的差异可从护士的评定结果中看出。记住，护士对老人进行评定时对实验一无所知，这样做是为防止她们在评定时产生偏见。他们证实，责任感提升组的被试在三个星期的实验过程后，其情况有显著的提高。与此同时，可以看出对照组的总体情况有所下降。事实上，“实验组中93%的被试（一人除外）的状况都得到了提高，而对照组却只有21%的被试（6名被试）的状况向积极方面变化”（P.196）。四楼的老人与别人的接触增多，与各类工作人员长时间地交谈。另一方面，责任感提升组的被试几乎很少从事一些被动性的活动，如仅在一旁观看工作人员的工作。

表5—3 问卷调查结果一览表

第一次管理和第二次管理的差异

问卷条目	责任感提升组	对照组	差异是否显著
被试的自我报告			
·愉快	+0.28	-0.12	是
·积极	+0.20	-1.28	是
·访谈者对被试机敏程度的评定	+0.29	-0.37	是
护士的评定			
·总体上的进步	+3.97	-2.39	是
·时间安排：			
—拜访其他老人	+6.78	-3.30	是
—拜访其他人	+2.14	-4.16	是
—和工作人员谈话	+8.21	+1.61	是
—观察工作人员的工作	-2.14	+4.64	是

（引自 P.195）

对行为的测量结果，进一步支持了控制力的积极作用。很明显，实验组中有较多被试出席了电影观摩。而在此前一个月，研究人员也放映过一场电影，但当时这种差异却并未显现出来。此外，虽然猜果冻豆比赛对于科学研究来说，似乎有点可笑，然而我们所得到的结果却非常有趣：有10位四楼的老人参与了这项游戏，但是二楼却只有1位老人参与了该游戏。

讨论

根据他们的研究，并结合其他人的早期研究，兰格和罗丁指出，对于一个被迫放弃控制力与自我决策权的人，如果我们给他一种较强的自我责任感，那么他的生活质量便会提高，生活态度会变得越发积极。对于这项研究的实际应用价值，作者简洁明了地指出：

为了改变那些真正减少老年人责任感或使老年人感到责任感降低的环境因素，我们可以建立也应该建立一些机制。而且，这项研究以及大量的文献资料显示，衰老和机敏性的降低不总是因年老而导致的不可避免的结果。事实上，研究表明，通过让老年人重新获得决策权和胜任感，一些因年老而引起的消极因素是可以被延缓、逆转或者防止的（P.197）。

研究发现的意义和后续研究

这些研究者在同一个疗养院，用同样的被试所做的后续研究也许能为这项研究结果的重要性提供最好的证明（Rodin & Langer，1977）。第一项研究结束的18个月之后，兰格和罗丁又返回阿登屋疗养院作进一步的跟踪调查，以了解责任感提升条件是否存在长期效应。对于依旧住

在该养老院的老人，由养老院的医生和护士对他们进行评估。此外，该文的作者之一（J.罗丁）与这些老人作了一次有关心理学和衰老的特殊讲座。对参加这次讲座的上次实验的被试人数进行了记录，并记录了提问的问题类型和频率。

护士对老人们的评估结果显示，责任感提升组的情况仍明显优于对照组。实验组的总平均数（把所有评定结果加起来，再计算平均分）是352.33，而对照组是262.00（统计差异极为显著）。从医生那里得到的健康评定值也表明，实验组的整体健康状况有所上升，而控制组则有轻微的下降。虽然在参加讲座的人数方面，两组没有显著差异，但大多数问题是由责任感提升组提出的，且其内容大多与自由和独立有关。在所有的研究结果中，最重要的发现也许是，在这18个月的间隔当中，有30%的对照组被试离开了人世，而在实验组中这个比例仅为15%。

对本项研究的一个重要批评意见，是由兰格和罗丁他们自己提出的。无论在何种情境下，只要涉及被试的身心健康，研究者都应从伦理道德的角度，十分慎重地对待干预的后果。我们在实验中使老年人重新获得了各类控制力，而在实验结束后，他们又失去了这些能力，这是危险而不道德的。在舒尔兹（Schulz，1976）的一项研究中，研究者让疗养院中的老人在大学生采访期间拥有程度水平不等的自主权。那些自主水平最高的人，无论在什么时候接受访问，无论被访问多久，都明显地表现出功能上的提高，这与兰格和罗丁的发现完全一致。然而，研究结束后，大学生们不再继续拜访老人（这里的部分原因是由于研究人员的疏忽）。这就导致了实验组比那些从未经历过控制力增加的被试衰老得更快。在兰格和罗丁的研究中，这个现象没有出现，因为老人对日常生活决策的总体控制感已经培养出来了，所以这一积极的变化能持续产生积极的结果。

近期应用

正如我们在前面所提到的，在人丰富多彩的一生中，个人的力量和控制力是一个非常关键的因素。这种潜力丧失之后，人也就进入了老年。兰格和罗丁的研究以及朱迪斯·罗丁（参见Rodin，1986）的一系列后续研究都清楚地说明了，人所拥有的控制力越多，其老年化过程就越快乐、越健康、越平稳。今天，养老院、国家养老机构认证部、医院以及其他一些机构都意识到了这一点，他们对老年人不断给予鼓励，并增加他们进行自我选择的机会、提高个人的力量和控制力。

包括兰格和罗丁1976年的研究在内的很多研究都支持这一观点，即当我们步入老年时，控制力对我们而言是必需的且有价值的。例如，2003年的一项关于德国老人院中的老人抑郁情况的研究发现，觉知到的自由和个人选择的减少，就像身体健康状况不良和缺乏社会支持一样，可作为抑郁症状的预测指标（Krampe等，2003）。作者总结道：“对老

年人抑郁的治疗和预防必须与有意的自我调节（个人选择）和可觉知的自由相结合。因为这二者在对抑郁的解释上互相独立。”

兰格和罗丁早期研究中所提出的问题被后人继续探索着。2010年，一项名为“生而选择：控制感需求的起源和意义”的研究提到，人类对周遭环境的控制能力、对行为结果进行个人选择的感知能力，不仅对心理健康有着重要价值，而且有可能是一种生物本能（Leotti，& Ochsner，2010）

这种生物意义上的必要性，似乎与选择及有意识地进行决策的能力有关。越来越多的来自人和动物的研究及脑扫描的研究模型表明，现代人类的控制感是从数百万年的进化过程中发展而来的一种生存机制。换句话说，能够做出选择、并对环境进行控制的早期人类更有希望生存下来，并将这种“控制基因”传递给下一代。这种基因传递的模式直到今天依然在继续。研究发现，当人们失去选择的机会和控制力时，就会产生压力。举个极端的例子，想象一下你的身体被限制活动了（例如被按倒或被绑起来），虽然你很清楚这些限制不会真的伤害你，但这种限制仍会导致强烈的应激反应（比如，心率加快、血压升高、肾上腺素释放）。为了更好地证明丧失控制感所带来的反应是“天生的”，研究者将同样的限制施加于动物，结果发现其反应模式与人类相似。

结论

可见，个人力量和控制力不仅影响着人的愉快心理，而且还影响着人的健康。将兰格和罗丁的思想运用于你的现实生活并不困难，请回想一下在哪些事件、情境和经历中，你对自己的行为几乎失去控制。你可能还记得与那些你能选择和操纵的事件相比，这些经历使你感到更加不安、更加生气、更加不愉快。在绝大多数的生活情境中，提高个人行为的选择程度的确是一个值得追求的人生目标。

Glass, C., & Singer, J.(1972).Urban stress: Experiments on noise and social stressors.New York: Academic Press.

Kramper, H., Hautzinger, M., Ehrenreich, H., & Kroner-Herwig, B.(2003).Depression among elderly living in senior citizen homes: Investigation of a multifactorial model of depression.Zeitschrift fur klinische psychologie und psychotherapie, 32，117-128.

Leotti, L., Iyengar, S., & Ochsner, K., (2010).Born to choice: The origins and value of the need for control, Trends in Cognitive Sciences，14, 457-463.

Rodin, J.(1986).Aging and health: Effects of the sense of

control.Science, 233 , 1271-1276.

Rodin, J., & Langer, E.J.(1977).Long-term effects of a control relevant intervention with the institutionalized aged.Journal of Personality and Social Psychology, 35, 897-902.

Rodin, J., Solomon, S., & Metcalf, J.(1979).Role of control in mediating perceptions of density.Journal of Personality and Social Psychology, 36, 988-999.

Schulz, R.(1976).Effects of control and predictability on the psychological well-being of the Institutionalized aged.Journal of Personality and Social Psychology, 33, 563-573.

第六章 情绪和动机

本章所关注的话题是我们内心的动机和情绪体验。很多非心理学家对用科学的方法探索这些问题感到困难重重。人们普遍认为，情绪和动机只是偶尔发生的，我们不能对其进行完全的控制，而且它们是与生俱来的一种本能。然而，心理学家们一直对情绪的产生以及人的行为动机怀有浓厚的兴趣。情绪和动机对行为的产生具有强有力的影响，而且在这一领域已有大量的研究成果，它们帮助我们更深刻地理解情绪和动机这两种心理现象。

本章的第一项研究也许会令人感到惊讶，这是马斯特斯（Masters）和约翰逊（Johnson）在20世纪60年代开展的有关人类性反应的研究。该研究之所以被收录在此，是因为人类的性感受和性行为受到情绪的强烈影响，而且它们本身也被视为是强有力的动机力量；第二项研究极负盛名且引人注目，它考察情绪的外部表达——面部表情，并得出了人类所拥有的基本面部表情具有跨文化的一致性结论；第三项研究探讨了极端情绪或所谓的“应激状态”对人体健康的影响；第四项研究使我们有机会领略动机研究领域最著名的实验之一，它首次证明了一个被称为“认知失调”的心理现象的存在。

研究21 性动机

Masters, W.H., & Johnson, V.E.(1966).Human sexual response.Boston: Little, Brown.

你或许并不认同人类的性行为和心理有密切关系的说法。很多人可能从逻辑上把性行为的研究划入生物学或生理学的学科范畴。这些学科的确在很多方面与性行为研究有关，并且主要是研究动物的性行为。然而，就人类而言，性毫无疑问是一种心理事件。仔细想想，性吸引、性欲以及性功能等都主要取决于人的心理。如果你对此存有疑义，请考虑一下以下几个显而易见的事实：你知道大多数人发生性行为的原因不仅仅是为了繁衍后代，更多的是心理方面的。此外，人类是地球上唯一受性问题困扰的物种，这些问题有性欲的压抑、性高潮缺乏、阳痿、早泄、阴道痉挛等。它们通常都涉及心理方面的原因。

但是，在讨论该项研究之前，我们应该知道，行为带来的欢愉以及对性问题的成功治疗离不开我们对性行为运作方式的清晰而完整的认识，即与人类性反应有关的生理学知识。这正是马斯特斯和约翰逊准备研究的问题。

在20世纪60年代以前，对人类性行为最具权威的研究工作当数20世纪40年代末、50年代初由阿尔弗雷德·金西（Alfred Kinsey）对美国人所做的大规模调查。著名的《金西性学报告》（Kinsey Reports），包括《男性性行为》（Sexual Behavcior in the Human Male）（1948）及《女性性行为》（Sexual Behavior in the Human Female）（1953），向数千名美国男性和女性询问了他们的性行为和性态度，内容涉及性交频率、手淫习惯以及同性恋经历，等等。随着这些调查报告的陆续发表，我们忽然有了一种衡量自己性生活方式的标准。金西的研究报告为我们提供了一个探索人类性行为的独特视角，直至今天，这些数据仍作为性行为的统计信息而为人们广泛引用。然而，虽然金西的研究是具有重大意义的，但是他的研究仅仅提供了人类性行为的内容，而在性行为的过程细节以及性困扰的处理方面留下了很大的空白。

于是，马斯特斯和约翰逊这两个名字已经成为人类性行为研究的同义词，并且为世界上数以万计的人们所熟知。20世纪60年代初，美国发生了一场现在被称为“性解放”的革命运动。剧烈的社会变革使人们能公开、坦率地对以前讳莫如深的性行为进行科学探索。直到20世纪60年代，根深蒂固的维多利亚式的价值观使人们对性遮遮掩掩，它成了不宜公开讨论的隐秘话题，更不用说加以研究了。所有这些都会使马斯特斯和约翰逊研究计划失去所有的社会支持和经济支持。但随着人们开始公开承认自己是具有性感受和性欲望的个体，社会不仅认可马斯特斯和约翰逊对此问题的研究，而且也需要这种研究。因此，原有的统计资料已

不能满足人们的需要。人们开始乐于了解自己面对性刺激时的生理反应。

正是在这样的社会背景下，马斯特斯和约翰逊始研究人类的性反应。他们的早期研究正是本文讨论的主题。尽管他们的研究工作是在30年前进行的，但它现在仍然是有关人类性反应的生理学知识的基础。

理论假设

马斯特斯和约翰逊研究的最重要假设是：为了理解人类的性活动，我们必须研究人们对性刺激做出的真实反应，而不仅仅是人们自己报告或以为的性体验。

他们指出，该理论的目的之一是为了心理治疗，即帮助人们克服他们可能正在经历的性问题。马斯特斯和约翰逊对他们的这一目标做了如下描述：

> 有关人类性行为的理论基础，只有在解答了下面两个问题之后，才能真正建立起来。这两个问题是：当男性和女性对有效的性刺激作出反应时，他们的身体发生了什么样的反应？当男性和女性对有效的性刺激作出反应时，为什么他们会作出这样的反应？如果医学工作者要对性功能缺陷者进行治疗的话，他们必须对这些基本问题做出令人满意的解答（P.4）。

在这一研究目的的基础上，马斯特斯和约翰逊还指出，能够获得上述问题答案的唯一方法便是对男性和女性性反应的各个阶段进行直接、系统的观察和生理测量。

方法

被试

正如你想象的，这样一个研究课题所遇到的第一个困难就是寻找被试。该研究要求被试愿意在实验室里从事性活动，同时接受严密的观察和监控。研究者担心这样的要求会导致根本找不到愿意参加研究的被试，而且很难保证所找到的被试在统计学上能够代表人类这一总体。研究者担心的另一点是，陌生的实验室环境可能使自愿参加该研究的被试不能以他们通常的方式作出反应。

在他们研究的早期阶段，马斯特斯和约翰逊雇用卖淫者作为被试。做出这一决定的原因是由于他们推测来自于一般的和典型的生活方式的个体会拒绝参加该项研究。8名女性和3名男性卖淫者接受了近两年的全面研究。研究者们认为，从这最初的11名被试身上所获得的研究数据，

对完善整个研究的研究方法和技术起了关键性的作用。

但是，这些被试对人类性反应的整体研究而言并不适宜。这是因为他们的生活方式和性经历与普通人还是相去甚远的。所以，研究者们深知，从这些被试身上所获得的任何结果不能确定无疑地推论到一般人群上去。因此，有必要获得一个更具代表性的研究样本。和他们最初的预测相反，研究者发现，寻找具有代表性的被试并不像他们先前想象的那么难。

通过与某大城市一个地区的一些学术、医疗机构的联系，马斯特斯和约翰逊找到了大量具有不同社会经济地位和教育水平的志愿者来参加这项研究。表6—1概括了最终被选为被试的志愿者在年龄、性别以及受教育水平等的人口学特征上的分布情况。所有的志愿者都接受了详细的访谈，以确定他们参加该研究的动机以及在性反应问题上与研究者沟通的能力。这些候选被试也都同意参加一次体检，以确定他们生殖能力的常态指标。

表6—1 被试的年龄、性别和文化程度等特征的分布情况

年龄	男性	女性	高中	大学	研究生
18-20	2	0	2	0	0
21-30	182	120	86	132	84
31-40	137	111	72	98	78
41-50	27	42	18	29	22
51-60	23	19	15	15	12
61-70	8	14	7	11	4
71-80	3	4	3	3	1
81-90	0	2	0	2	0
总数	382	312	203	290	201

（引自 P.13-15）

程序

为了仔细地研究人类性活动和性刺激过程中的生理反应，研究人员有必要设计出极为精密的观察测量方法，其中包括对脉搏、血压、呼吸频率等生理反应的标准化测量。此外，还要观察和记录特定的性反应。为此，“被试在各个时段的性活动将包括各种人工操纵和机械操纵，男性被试与其女伴以仰卧、跪式或最习惯的姿势进行自然的性交，女性被试则以仰卧或跪着的姿势进行人工操纵的性交”（P.21）。这就意味着，研究者不仅观察和测量被试以各种姿势进行的自然性交，而且还要观察和测量在人工或机械装置的帮助下所进行的手淫活动，这些机械装置是为对被试的反应进行清晰的记录而专门设计的。

这些由物理学家设计的特殊装置主要是由塑料制成的透明人造阴茎，这些人造阴茎可以帮助研究者进行真切的内部观察。人造阴茎可以根据女性感觉舒适与否来调整大小，而且在整个反应周期内完全由女性

被试控制其在阴道中插入的深度和活动频率。

被试的适应和舒适感

不难想象，所有这些期望、观察和装置都可能会使很多被试产生实际的情绪困扰。马斯特斯和约翰逊也的确意识到了这一点。为了帮助被试在整个研究过程中始终处于舒适的状态，他们保证：

> 在研究开始之前，研究者先在私下里积极鼓励被试的性活动，然后被试的性活动在调查组成员在场的情况下继续进行，直到被试在实验环境中泰然处之。……只有当被试在他们的环境中感到很安全，而且相信自己能够完成研究任务的时候，研究人员才对他们的反应进行记录。这一训练阶段是为让被试对研究的单纯性和研究承诺的匿名方式建立信任感（P.22—23）。

有些被试只参与了一段时间的研究便退出了实验，而另一些被试则积极参加了长达数年的研究。在这项研究中，马斯特斯和约翰逊估计他们共观测了约10000个完整的性反应周期，其中女性和男性的数量之比是3：1。换言之，“至少有7500个完整的性反应周期是由女性被试在各种研究程序中完成的。相反，男性被试的性高潮（射精）经历的研究数据至多仅为2500个”（P.15）。

结果

马斯特斯和约翰逊从该研究中发现了与人类性反应有关的极为有价值的信息。稍后我们将对这些发现的精华部分进行概括。不过，另一个值得人们牢记的方面是，他们从被试那里所得到的研究数据几乎适用于每一个人。当然，在人类整体中总会存在一些罕见的例外、异常或变态反应，但除了这些极个别的现象外，每个人对性刺激的生理反应都是极为相似的。这一事实致使马斯特斯和约翰逊的这一研究的价值无法估量。它为改善每个人的性生活提供了可能。当你读到他们早期的发现时，不应忘记，他们并未研究与人类性活动有关的态度、情绪、价值观、道德、偏好、适应性或喜恶。很明显，这些都是因人而异的，而且这种个体差异造就了人类性行为的极大的、甚至是不可思议的多样性。那就让我们见识一下马斯特斯和约翰逊的这些最具影响力的研究发现吧！

性反应周期

在研究了大约10000个性行为案例后，马斯特斯和约翰逊发现，人类的性反应可以分为四个阶段，并将其命名为“人类性反应周期”。这些阶段依次为兴奋期、持续期、高潮期和消退期（参见表6—2）。虽然两

位心理学家在自己的著作中承认这些阶段是人为划分的，但这种划分使得对性反应的讨论变得更加容易和清晰。如今，学术界或专业领域在对人类性反应进行讨论时，无不参考这四个反应阶段。尽管男性和女性在性反应上存在一些差异（这一点我们将稍后加以讨论），但男性和女性在性行为中都要经历相同的四个阶段。

表6—2 性反应周期的四个阶段

阶段	女性的反应	男性的反应
兴奋期	第一标志：阴道润滑。阴蒂头勃起。乳头变硬，乳房增大。阴道变长且阴道靠里的三分之二的部分有所扩展	第一标志：阴茎勃起。（阴茎勃起的时间因个体差异、年龄、酒精或药物的使用、疲劳和应激等状态的不同而有所不同）。阴囊和睾丸提起。如果注意力转移，勃起的阴茎可能会缩回，但也很容易恢复到勃起状态
持续期	阴道靠外部的三分之一膨胀起来，阴道口缩小至原来的50%。阴道靠里的三分之二继续膨胀或“充塞”。阴蒂缩回。润滑程度降低。小阴唇充血且颜色变深，预示着性高潮将至。肌肉紧张，血压上升	阴茎完全勃起。如果注意力转移，阴茎也不会轻易地缩回，龟头进一步变大。尿道球腺分泌精液。睾丸进一步上升、旋转并增大，预示着性高潮将至。肌肉紧张，血压上升
高潮期	骨盆开始以0.8秒的时间间隔有节奏地收缩。阴道内下壁的这种收缩感觉尤其明显。子宫也有节奏地收缩。全身肌肉紧张。记录到的持续时间从7.4秒到104.6秒不等。时间长度并不等同于可感觉到的肌肉紧张度，即不是持续时间越长，肌肉紧张的强度越大	骨盆开始以0.8秒的时间间隔收缩。射精，即精液的射出分为两个阶段：（1）准备发射：精液聚积在尿道口处，感到射精已无法避免；（2）排出：阴茎的肌肉收缩，迫使精液从尿道排出
消退期	阴蒂、子宫、阴道和乳头等在一分钟之内恢复到原来未唤醒的状态。在5到10分钟之内，阴蒂通常仍对触摸非常敏感。如果女性尚未体验到性高潮，那么这一消退过程可能会持续几个小时	大约有50%的男性勃起的阴茎在一分钟内会缩回；随后逐步地完全恢复到未唤醒的状态。睾丸变小并下降。阴囊松弛下来

性解剖学

马斯特斯和约翰逊对该项研究的伟大贡献之一便是揭开了人类性行为的神秘面纱。研究者试图纠正在性的解剖学知识方面存在的普遍误解，特别是对阴茎和阴道的分析。有史以来，男性最常见的性担忧便是阴茎的大小问题。马斯特斯和约翰逊研究了很多阴茎，且最终找到了该问题的科学答案。他们称这种困扰为“性器误解”（phallic fallacies）。男性通常为这两个问题而苦恼：（1）阴茎越大才越能满足女性的需求；（2）他们怀疑自己的阴茎太小。马斯特斯和约翰逊对他们在研究中测量到的阴茎平均大小进行了讨论，并表明上述想法是被误导的结果，而且解释了阴茎和阴道在两性性交过程中的机能。

研究者发现，在80名男性被试中，松弛状态下，阴茎长度的正常范

围是从2.8英寸到4.3英寸，其平均长度大约为3英寸。在勃起状态下，阴茎长度的正常范围是从5.5英寸到近7英寸，其平均长度大约为6英寸。这些数字明显小于人们传统观念中所认为的大阴茎和小阴茎的长度。但更令人惊奇的是，当研究者测量勃起阴茎的长度时，他们发现松弛状态下较大的阴茎并不预示着它在勃起时也更大些。事实上，总体而言，他们发现松弛状态下较小的阴茎在性兴奋时比松弛状态下较大的阴茎有更大的变化幅度。看一看下面的平均数值，松弛状态下3英寸长的阴茎可以增加到6英寸，而松弛状态下4英寸长的阴茎只增加了大约2.5英寸并最终达到6.5英寸。为了进一步说明这一发现，马斯特斯和约翰逊报告了所观察到的在松弛状态和勃起状态下变化幅度最大和最小的阴茎。一位被试的阴茎在松弛状态下长度是2.8英寸，其勃起状态下的阴茎长度增加了3.3英寸，总长度达到了6.1英寸。另一名被试的阴茎在松弛状态下是4英寸，勃起时仅增加了2.1英寸，其最终长度同样也是6.1英寸。

比以上这些阴茎的测量更重要的是女性的性快感和性满足取决于阴茎大小的观念。正像他们书中的“对阴道的误解”一节所说明的那样，研究发现，以上这种观念是毫无根据的。在使用上述人造阴茎技术所做的观察中，他们得出结论，阴道是一种足以容纳大小不同阴茎的极富弹性的器官。“不管阴茎的大小如何，阴道总能在阴茎最初的几次抽动中使阴茎充满阴道”（P.194）。此外，他们发现，在性反应周期的持续阶段（见表6—2），阴道口的内壁膨胀起来，足以紧紧缠绕住任何大小的阴茎。因此，作者的结论是“显而易见，阴茎的大小在对女性的性刺激上通常只是个无足轻重的因素”（P.195）。

性反应上的两性差异

尽管男性和女性在性反应周期上存在很多相似之处，但马斯特斯和约翰逊也发现了一些重要的差异。他们发现的最著名且最具革新意义的差异涉及性反应周期的高潮阶段和消退阶段。高潮阶段过后，男性和女性都进入消退阶段，此时性紧张迅速降低，性器官还原为以前未唤醒的状态（这就是所谓的消肿）。马斯特斯和约翰逊发现，此时男性存在一个不应期，即在此期间，不管他所受到的性刺激的类型或数量如何，他在身体上都不能经历另一次性高潮。这一不应期可以持续几分钟，几个小时，或者甚至是一天，而且随着年龄的增长，不应期所持续的时间有不断延长的趋势。

马斯特斯和约翰逊发现女性并没有明显的不应期，而且在有效的刺激下，女性能够在第一次性高潮后紧接着经历一次或多次性高潮，且不失去任何性唤醒的能力。和男性的情况不同，在性交过程中女性“能够使自己的性高潮体验持续相对更长的一段时间”（P.131）。

尽管对很多女性而言，这种多次性高潮的能力已不是什么新闻，但这并不广为人知。在马斯特斯和约翰逊的研究以前，人们普遍相信男性有更强的性高潮能力。因此，这一发现与马斯特斯和约翰逊的其他很多

发现对人们所持有的男女性行为方面的文化和社会态度产生了深远的影响。值得注意的是，尽管女性在生理上能够体验多次性高潮，但并非所有的女性都会去寻求或甚至渴望多次性高潮。事实上，很多女性从未体验过多次性高潮，但也非常满意自己的性生活。此外，很多体验过多次性高潮的女性称，一次性高潮就能使她们感到满足。就性行为的生理和情绪方面的满意度而言，个体之间的差异是相当大的。马斯特斯和约翰逊试图继续探讨性行为的生理反应的全部可能性。

批评

对马斯特斯和约翰逊早期研究的大部分批评主要是：（1）他们对性反应四个阶段的划分过于武断；（2）他们几乎没有讨论性行为的认知和情绪方面的内容。马斯特斯和约翰逊在他们的早期著作中对这些批判也都有过回应。

如前所述，两位作者已充分意识到他们对反应周期的划分纯粹是人为的，但这些划分有助于人们研究和解释人类性反应的复杂过程。多年来，其他的研究者提出了不同的阶段理论。例如，海伦·辛格·卡普兰（Helen Singer Kaplan，1974）提出了一个包括欲望、血管充血（生殖器的饥渴状态）和肌肉收缩（性高潮）的三阶段模型。这些阶段反映了卡普兰的信念，即在任何性刺激开始之前，对性反应的分析应该始于性欲的产生，而且她认为在兴奋期和持续期之间没有必要做明确的区分。她对性欲的强调引发了对马斯特斯和约翰逊早期研究的另一种主要批评意见，即他们的研究缺乏对性反应的心理因素的关注。

然而，马斯特斯和约翰逊再次明确指出，检测性行为的心理因素并不是他们此项研究的目标。不过他们相信，全面了解与性行为有关的生理知识对完满的性生活而言，绝对是一个先决条件。而且，他们后来还出版了涉及人类性行为的心理和情绪方面的著作，并在其中对这个问题作出了令人信服的论述。

最后，自马斯特斯和约翰逊的第一部著作问世以来的三十多年间，一直有研究者对他们把这些研究发现推论到所有人身上的可行性持有怀疑。例如，有研究发现一小部分女性也有不应期，在此期间，她们不能体验性高潮，此外也有一小部分男性能够体验多次性高潮，其间几乎不存在不应期。此外，虽然射精一向被认为是男性的专利，但最近研究发现，在某些情况下，一些女性性高潮时也会射精（参见Zaviacic于2002年对该研究的讨论）。

近期应用

在此，要想从每年参考马斯特斯和约翰逊早期研究的出版物中寻找

具代表性的一部分，并将其罗列出来，这似乎是不可能的。这些出版物涉及的范围很广，从有关性的核心教科书（参见Hock，2007；McAnulty & Burnette，2004）到在心理学和性行为杂志中处于前沿的文章。

此外，如你所料，马斯特斯和约翰逊的模型过去和现在都饱受争议。目前，主要的争论点仍然围绕着性反应四阶段模型是否如研究者所认为的那样，适用于所有男性和女性。

有一项研究结合了基马斯特斯和约翰逊的开创性研究，针对女性性生活满意度，在全国范围内实施了问卷调查，被调查者为近1000名年龄在20~65岁之间的异性恋女性（Bancroft等，2003）。该研究的目的是调查女性性问题是否被视为与男性的相似，以及什么药物治疗能对女性有效。在这一方面，一些治疗勃起障碍的药物（例如伟哥、艾力达、西力士）已经对男性有明显疗效。该研究发现性反应的生理表现（唤起、阴道润滑、高潮）与被调查者的性烦恼之间并没有较高的相关："总体上讲，与女性性生理反应障碍相比，在性交过程中缺乏幸福感和负性情绪体验会对女性的性困扰起到更加决定性的作用。虽然我们没有男性在此方面的直接对应数据，但我们可以预测该模式会有不同，对男性而言，生殖器的反应会更加重要。"（Bancroft et al.，2003，P.202）换言之，女性最普遍的性问题的解决方法远比一粒"粉色药片"复杂得多。

实际上，在2000年，由12名女科学家、研究人员和临床医生组成的研究队伍采用新的研究方法来探讨女性的性问题。她们认为男性和女性之间在性问题上不同之处多于相似之处，同时认为马斯特斯和约翰逊的四阶段模型在描述、解释和治疗女性的性问题上是无效的（参考Tiefer，2001）。这种关于女性性问题的新观点声称："女性的阶段数并不恰好等同于马斯特斯和约翰逊模型。比如女性通常不能区分欲望和刺激，女性对生理刺激的关注比对主观唤醒要少。"（Tiefer，2001，P.93）研究人员指出马斯特斯和约翰逊的模型在很多方面将男性和女性的性反应等同看待，从而忽略了理解女性的性问题必需的重要因素。这其中包括性反应发生时间和女性的性反应模式的个体差异之间的关系。更独特的是，他们认为女性面临的性困扰需要建立一个分类系统，这个系统应包含文化、政治和经济背景（比如性教育或避孕措施的缺乏）、女性的伴侣和关系问题（如对虐待的恐惧、力量失衡和整体的不协调）、心理因素（比如以往的性创伤经历、抑郁和焦虑）以及医学因素（如激素失衡、性传播疾病和药物的副作用）等。

感谢马斯特斯和约翰逊所做的大量工作，使我们对包括人类性愉悦和性反应在内的生理过程的理解与前半个世纪相比，前进了一大步，但是仍有大量问题有待研究。毋庸置疑，在以马斯特斯和约翰逊开创性的工作为背景下，相关研究会继续进行，我们关于人类性反应的视野也会逐渐变宽。

结论

马斯特斯和约翰逊于1971年步入结婚殿堂。在接下来的二十多年里，他们作为一个小组，继续进行相关工作和发表文章。在1992年，由于在研究方向的分歧日趋增加及退休的原因，这对夫妇离婚了，约翰逊也随之退休了。马斯特斯一直担任着圣路易马斯特斯和约翰逊研究院的院长直到1994年退休。他于2001年2月16日死于帕金森症，享年85岁。

让我们回到本讨论的开头部分，马斯特斯和约翰逊研究的主要目的是要探讨性功能不全的问题，以便帮助人们解决它们。毫无疑问，他们达到了这一研究目的。几乎所有的性治疗，从勃起问题、性高潮缺乏到早泄、性唤醒困难等，都是建立在马斯特斯和约翰逊的研究的基础上的。马斯特斯和约翰逊为我们理解和研究人类性行为作出了难以估量的贡献。最近，人们对有关性知识的教科书进行调查后发现，和其他的研究者相比，马斯特斯和约翰逊的研究被引用的次数更多，所占篇幅也更多。除此之外，在《人类的性反应》（Human Scxual Response）一书出版后的数十年间，马斯特斯和约翰逊继续着他们的研究工作，并将其研究发现应用到帮助人们实现性生活和谐的实践当中。在他们的处女作问世4年后，他们又出版了《人类的性机能障碍》（Human Sexual Inadequacy）一书（1970），该书包括他们的早期研究，且直接将其成果应用于解决性问题上。我们可以从他们随后出版的一系列著作中发现他们不断前进的脚印，这些著作包括：《快乐的结合》（The Pleasure Bond，1970）、《同性恋观》（Homosexuality in Perspective，1979）、《人类的性行为》（Human Sexuality, 5/e，1995）、《危机：艾滋病时代的异性性行为》（Crisis: Heterosexual Behavior in the Age of AIDS,1988）、《马斯特斯和约翰逊论性与人类的爱情》（Masters and Johnson on Sex and Human Loving，1986）、《异性恋》（Heterosexuality，1998）。

Bancroft, J., Loftus J., & Long J.(2003).Distress about sex: A national survey of women in Heterosexual relationships.Archives of Sexual Behavior, 32，193-208.

Hock, R.R.(2007).Human sexuality.Upper Saddle River, NJ: Pearson Prentice Hall.

Kaplan, H.S.(1974).The new sex therapy.New York: Brunner/Mazel.

Kinsey, A., Pomeroy, W., Martin, C., & Gebhard, P.(1948).Sexual behavior in the human male.Philadelphia: W.B.Saunders.

Kinsey, A., Pomeroy, W., Martin, C., & Gebhard, P.(1953).Sexual behavior in the human female.Philadelphia: W.B.Saunders.

McAnulty, R.D., & Burnette, M.M.(2004) Exploring human sexuality: Making healthy decisions, 2nded, Boston: Pearson Allyn & Bacon.

Tiefer, L.(2001).A new view of women's sexual problems: why new? Why now? Journal of Sex Research, 38, 89-96.

Zaviacic, Milan (2002).Female urethral expulsions evoked by local digital stimulation of the g-spot: Differences in the response patterns.The Journal of Sex Research, 24, 311-18

研究22 我能读懂你的面部表情

Ekman, P., & Friesen, W.V.(1971).Constants across cultures in the face and emotion.Journal of Personality and Social Psychology，17，124-129.

如果你想到一些有趣的事，你的脸上会浮现出什么表情？再回想一下过去的伤感经历。你的面部表情会发生变化吗？答案是肯定的。毫无疑问，你能意识到特定的面部表情是与特定的情绪相对应的，而且在大多数情况下，你能通过某人的面部表情识别出他此时此刻的情绪情感。现在设想一下，如果一个人来自另一种文化背景，如罗马尼亚、苏门答腊，抑或蒙古，你还能借助他们的面部表情成功地识别他们的情绪状态吗？换句话说，你是否相信表现情绪的面部表情具有普遍性？大多数人相信他们能够辨认来自异域文化的人们的面部表情，但当他们考虑到异域文化与他们自身的文化差异悬殊时，又会变得犹豫不决起来。试想，不同文化在动作手势、个人空间、礼仪规范、宗教信仰以及态度等方面是多么不同。所有这些差异都会影响人的行为，因此，如果有任何人类特征（包括面部表情）在所有文化中是完全一样的话，那该是多么不可思议啊！

在面部表情研究领域，保罗·埃克曼（Paul Ekman）是站在最前沿的卓越研究者。他的这篇早期论文详细描述了他的研究工作，该研究旨在证明面部表情具有普遍性。尽管作者在论文的引言部分承认，先前的研究者已经发现了一些证据支持人的面部表情受到文化差异的影响，但他们认为这些证据是不充分的，不足以推翻他们的论断，即表达基本情绪的面部表情具有跨文化的一致性。

在开展该项研究的几年前，埃克曼和弗里森已经着手研究这方面的问题，他们向具有大学文化程度的阿根廷、巴西、智利、日本和美国被试呈现人脸部的照片。结果来自所有国家的被试都成功地识别出对应于某种情绪的面部表情。研究者把这一发现作为面部表情具有普遍性的证据。然而，正如埃克曼和弗里森他们自己所指出的，由于参加上述研究的被试都能够接触到像电影、杂志、电视等国际性的大众传媒，而这些传媒又充斥着面部表情，因此，他们欢迎各方面的人对他们的研究提出批评意见。只有找到一种从未接触过上述大众传媒的文化群体，才有可能证明面部表情具有跨文化的一致性。试想，在当今世界要想找到这样一种文化是多么困难（也许就根本不可能），即使在当时的1971年，要做到这点也很不容易。

埃克曼和弗里森来到了位于新几内亚东南部的高原为他们的研究寻找被试，在那里生活的弗尔族（Fore）人与世隔绝，其社会状态仍处于石器时代。很多当地居民几乎没有接触过西方或东方的现代文化。因

此，除了自己的面部表情外，他们还没有接触过来自其他文化的表现情绪的面部表情。

理论假设

埃克曼和弗里森的研究所依据的理论假设是，对应于基本情绪的特定面部表情具有跨文化的一致性。埃克曼和弗里森用极为简洁的话对此进行陈述：

> 本文的目的是考察处于前文明时期的人们，尽管他们没有接触过任何的异域文明，他们依然能和来自西方文明和东方文明的人们一样，从相同的面部表情中识别出相同的情绪概念（P.125）。

方法

南弗尔族（the South Fore）是当地最与世隔绝的弗尔族分支。研究者从中选择的被试从未看过电影，不会说英语或洋泾浜英语（一种掺杂了当地语言的英语），从未为西方人工作过，并且也不住在西方人在当地的居住地内。共有189名成人和130名儿童被选为被试，他们都来自总人口为11000人的南弗尔族。对照组则选取了通过看电影、上教会学校等方式接触过西方文化的23名成人。

经反复尝试，研究者发现，让被试最有效地识别情绪的方法是给他们呈现三张印有不同面部表情的照片，同时读一段简短描述该情绪产生的情景或故事情节给被试听。每一段描述对应于其中一张照片。听完故事后，要求被试指出与某一故事情节最为匹配的面部表情。在故事的选择上，主试非常谨慎小心以确保每种情境只与一种情绪有关，而且它必须能为南弗尔族人所辨认。表6—3列出了由埃克曼和弗里森编制的六个故事。作者解释说，与恐惧有关的故事必须编得长些，以避免被试将它与惊奇或愤怒相混淆。

研究者用男人、女人、男孩、女孩共24人的40张照片，作为六种情绪表现的样片。这些照片经此前来自其他文化背景的被试使用已被证明是有效的。每张照片在被作为某种情绪的代表之前，至少须经来自东西方两种文明的被试中70%的人的判定。

表6—3 埃克曼和弗里森的对应于六种情绪的故事

情绪	故事
1. 愉快	他的（她的）朋友来了，他（她）很高兴。
2. 悲哀	他的（她的）孩子（妈妈）死了，他（她）感到很悲伤。
3. 愤怒	他（她）很生气，而且要动手打架。
4. 惊奇	他（她）看到一件出乎意料的新玩意儿。
5. 厌恶	他（她）看见一件他（她）不喜欢的东西；或者他（她）闻到某种很难闻的东西。
6. 恐惧	他（她）独自坐在自己的屋子里，整个村子里没有其他人。屋子里没有刀、斧头或弓箭。一头野猪正站在他的（她的）家门口，这个男人（女人）正盯着这头野猪而且非常害怕它。这头野猪一直在门口站了好几分钟，这个人正很害怕地看着它。这头野猪没有离开的意思，他（她）害怕它会咬他（她）。

（引自 P.126）

在正式的实验中，由一名研究人员和一名充当翻译的当地人组成一个主试小组，由充当翻译的当地人向被试解释任务，并把故事翻译成当地的语言讲给他们听。主试向每一位成人被试出示三张照片（其中一张是正确答案，其他两张是错误答案），然后告诉被试与正确的那张照片相匹配的故事情节，最后要求被试挑出那张印有与故事相匹配的表情的照片。儿童被试的实验程序与成人被试的完全一样，有所不同的是，主试只要求他们从两张照片中挑出正确的答案。给每个被试提供的照片多组，并且确保每张照片不会在比较中重复出现。

充当翻译的当地人都经过严格的训练，以使他们不会对被试施加任何影响。主试告诉他们没有绝对正确的反应，要求他们不要催促被试回答。此外，还教他们如何每次以同样的方式翻译同一个故事，如何使自己不对故事进行扩充或润饰。为了避免无意识的偏见，研究小组的西方研究人员不能观察被试，只负责对被试的回答进行录音。

请记住，所提供的照片中有西方人的面部表情。那么，弗尔族人是否能在甚至没见过西方人的情况下正确识别照片中的表情所传达的情绪呢？

结果

首先，埃克曼和弗里森分析被试的回答，看是否存在性别差异以及成人和儿童间的差异。结果发现：成年妇女对参加实验表现得更加迟疑，她们比成年男子更缺少与西方人的接触机会。但不管怎样，在正确识别照片所代表的情绪的能力上既不存在性别差异，也不存在成人与儿童之间的差异。

表6—4和表6—5分别显示了成人和儿童在识别六种情绪时正确反应的百分比。需要指出的是，研究者并未要求所有的被试识别所有的情绪照片，而且有些被试可能会被要求重复识别同一种情绪照片。因此，表中被试的数目并不等于参加实验的被试总数。被试除了在辨别恐惧和惊

奇时的正确率稍低外，其他正确反应的百分比都达到了统计学上的显著水平。在辨别恐惧和惊奇时，出现了较多错误，甚至还有一组被试当故事情节描述的是恐惧情绪时，他们选择惊奇的次数竟占到了总次数的67%。

表6—4 成人正确识别照片所表现的情绪的百分比

故事所描述的情绪	被试数量	正确识别的百分比（%）
愉快	220	92.3
愤怒	98	85.3
悲哀	191	79.0
厌恶	101	83.0
惊奇	62	68.0
恐惧	184	80.5
恐惧（含有惊奇的成分）	153	42.7

（引自P.127）

表6—5 儿童正确识别照片所表现的情绪的百分比

故事所描述的情绪	被试数量	正确识别的百分比（%）
愉快	135	92.8
愤怒	69	85.3
悲哀	145	81.5
厌恶	46	86.5
惊奇	47	98.3
恐惧	64	93.3

（引自P.127）

将受西方文化影响的被试（即对照组被试）与未受西方文化影响的成人被试比较，他们在正确匹配照片与故事的比例上不存在显著差异。年龄稍大的儿童与年幼儿童之间也不存在显著差异。正如你在表6—5中所看到的，儿童的成绩似乎比成人要好，但埃克曼和弗里森把这一现象归因于儿童完成的是二选一的任务，而成人完成的则是三择一的任务。

讨论

根据所获数据，埃克曼和弗里森便毫不犹豫且非常自信地下结论说："成人和儿童的研究结果都明确地支持了我们的假设，即特定的面部行为与特定的情绪之间存在普遍的联系。"（P.128）这一结论的基础是南弗尔族人没有机会掌握任何与西方人面部表情有关的知识，这样，除非面部表情具有普遍性，否则他们是无法进行正确识别的。

为了对埃克曼的研究结果进行复核（double-checking），研究者用录像机录下了当地人演示的同样六种面部表情。后来，他们把这些录像带放给美国大学生看，并要求他们对表情进行识别，他们都正确地识别了对应于每种情绪的面部表情。

上述两项研究的研究结果都否定了“与情绪相关的面部表情具有文化特异性”的观点，这种观点认为面部表情是一种与文化紧密相联的独特传统习俗体系，它对另一种文化的成员而言是不可理解的（P.128）。

和上述发现不一致的一个例外是，被试在识别恐惧和惊奇这两种面部表情时出现了混淆现象。埃克曼和弗里森在解释这一例外时承认，情绪的面部表现确实存在某种文化差异，但这并不足以否认几乎其他所有的面部表情都能被来自不同文化的人们正确识别。研究者推测恐惧和惊奇的混淆“也许是因为在该文化中，恐惧的事件几乎总是令人惊奇的，例如，邻村的一名充满敌意的成员突然出现；一场意想不到的招魂术或者巫师的表演，等等”（P.129）。

研究影响

埃克曼和弗里森的研究科学地证实了人们对面部表情具有普遍性的猜测，然而，你或许仍会产生这样的疑问：“这一研究的意义何在？”关于这一问题的答案，我们也许可以在心理学有关人类行为的先天—后天争论（人类行为究竟是先天性的还是后天习得的）中找到。由于该研究中所涉及的六种情绪的面部表情几乎不受文化差异的影响，我们也许可以下结论说这些情绪应该是先天性的。也就是说，人在出生时这些情绪就已在生物学意义上形成了。

行为学家对于表情具有普遍性的观点颇感兴趣的另一个原因在于它已涉足了人类的进化问题。1872年，达尔文出版了他的经典著作《人类和动物的表情》（The Expression of Emotion in Man and Animals）。他认为，面部表情是一种适应机制，它有助于动物适应它们的环境进而提高其生存能力。从这一观点可以作进一步的推论，如果某些特定的信息能通过面部表情在动物的某个种系中或种系之间进行交流，那么，这将增加物种的生存概率。例如，恐惧的表情是一种无声的警告，它预示着食肉动物的出现将会带来的危险；愤怒的表情可对地位低下的同类成员发出警告，要求它们从自己的身边走开；厌恶的表情所传递的信息是：“你做什么都行，就是千万不要吃它！”从而避免了一次潜在的食物中毒。然而，如果这些表情在所有的个体间不存在普遍性，那么它们对动物将毫无作用。虽然这些面部表情对提高人类的生存能力已不那么重要，但这些表情具有普遍性的事实却表明，我们是从处于进化进程中的祖先那里继承了这些表情的，它们帮助我们达到了现在所处的进化高度。

另一项引人注目的研究证明了人类面部表情曾经具有的生存价值。研究者（Hansen & Hansen，1988）推测，如果某些诸如愤怒等的面部表情能够对迫近的危险起报警作用，那么它们相对于其他不具威胁性的面部表情（如愉快等）应更易被人们所辨别。为了检验这一假设，他们向被试出示一些照片，这些照片上的人们具有各不相同的面部表情。在某些照片中，除了一人的表情是愤怒的以外，其余人的表情都是愉快的；而另一些照片则相反，即只有一人表情是愉快的，其余人的表情都是愤怒的。被试的任务是找出照片中那个具有不同表情的人。被试从很多愤怒表情中寻找愉快表情所需的平均时间，显著长于从很多愉快表情中寻找愤怒表情。此外，随着照片中的人数的增加，被试寻找表情愉快的人所花费的时间也相应增加，但寻找愤怒表情的人的时间却没有显著增加。这一发现以及其他类似的研究发现都表明，人类也许天生就对这类能提供生存信息的表情比较敏感，因为这种表情所传递的信息与存活息息相关。

近期应用

来自其他领域的许多研究也借助埃克曼的早期发现，努力加深我们对存在发育或学习缺陷的儿童和成人的了解。其中一项研究发现，被诊断为自闭症（一种广泛性发育障碍，特征为语言缺陷、社会退缩和重复自我刺激的行为）的儿童对基本情绪的面孔识别存在困难（Bolte & Poustka，2003）。这种困难在不只有一个自闭症儿童的家庭中更明显，这也许可以解释为什么自闭症个体对其他人的情绪反应的解释存在困难。

然而，埃克曼研究的影响不仅限于对人类。他1971年的这项研究被一篇关于农场动物情绪的文章（Desire，Boissy，& Veissier，2002）引用。这些研究者认为，农场动物所产生的效益在一定程度上取决于它们对周围环境的情绪反应。当某一动物与周围环境和谐相处时，效益实现最大化；但是“如果个体感觉到任何显著的不和谐，都会导致由于负性情绪体验带来的效益降低”（P.165）。

另外一项引用了埃克曼1971年文章的研究致力于探讨特殊面部特征——眉毛——对面孔再认的影响（Sadr，Jarudi，& Sinha，2003）。先前的研究更多是对眼睛、嘴进行的，而这些研究者发现，眉毛可能比眼睛的作用更重要。作者提到“在熟悉面孔中缺少眉毛会导致显著的再认障碍”。实际上，缺少眉毛比缺少眼睛更能显著地引起面孔再认的困难（P.285）。所以，如果你需要有效的伪装，一定要把你的眉毛遮起来！

结论

自埃克曼完成表情的跨文化早期实验研究的三十年以来，埃克曼或独自，或与弗里森等其他研究者合作，继续进行他的研究工作。在这一系列的研究中，有很多引人注目的研究发现。埃克曼后来涉足的“表情的反馈理论”（facial feedback theory）便是一个例子。这一理论主张，你的面部表情实际上又把所有的表情信息反馈到大脑，以帮助你解释正在经历的情绪体验。埃克曼通过鉴别与上述六种基本情绪有关的面部肌肉活动来检验他的观点。他指导被试控制不同的面部肌肉来模仿各种情绪，在此过程中，埃克曼便能据此测量与特定情绪相对应的生理反应指标，这种特定情绪仅是由特定的面部表情所导致，而不是测量真正的情绪本身（Ekman，Levensen，& Friesen，1983）。

埃克曼还把自己的研究领域扩展到对欺骗的研究，他还探索面部表情与形体表情如何泄露说话者本人是否在撒谎的信息。他的发现表明，仅观察说话人的面部表情时，人们正确判断说话人在说谎的水平大约略微高于随机水平；然而，在主试允许被试观察某人的整个身体时，被试的正确判断率将大大提高。这表明，人的整个身体或许比单独的面部表情更能提供与某种心理状态相关的线索（参见1985年埃克曼对此问题所做的详细讨论）。近期，埃克曼将他宽泛的研究内容整合到名为《情感揭示：通过识别面部表情和情绪来提高沟通技巧与改善情感生活》（Emotions Revealed: Recognizing Faces & Feelings Improve Communication and Emotional Life）的著作中，该书可以使读者应用作者的成果来解读面部表情所表达的意思，从而帮助我们提高与的伴侣、孩子、同事甚至是陌生人的交流与互动（Ekman，2007）。

埃克曼和他的合作者已经在用面部表情进行的非言语交流的研究方面为我们提供了大量文献，并且他们在这方面的研究仍在继续。毋庸置疑，该领域的研究将持续到我们能够最终实现我们的研究目标，这一目标正如埃克曼和弗里森在1975年出版的一本书《摘下你的情绪面具》（Unmasking the Face）的书名所描绘的那样。

Bolte, S., & Poustka, F.(2003).The recognition of facial affect in autistic and schizophrenia subjects and their first-degree relatives.Psychological Medicine, 33, 907-915.

Darwin, C.R.(1872).The expression of the emotions in man and animals.London: John Murray.

Desire, L., Boissy, A., & Veissier, I.(2002).Emotions in farm animals: A new approach to animal welfare in applied ethology.Behavioral Process, 60，165-180.

Ekman, P.(1985).Telling lies.New York: Norton.

Ekman, P.(2003).Emotions revealed: Recognizing faces and feelings to improve communication and emotional life.New York:

Times Books.

Ekman, P.(2007).Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and emotional life.New York: Henry Holt.

Ekman, P., & Friesen, W.(1975).Unmasking the face.Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P., Levensen, R., & Friesen, W.(1983).Autonomic nervous system activity distinguishes between emotions.Science ，164, 86-88.

Hansen, C., & Hansen, R.(1988).Finding the face in the crowd: An anger superiority effect.Journal of Personality and Social Psychology, 54, 917-924.

Sadr, J., & Sinha, P.(2003).The role of eyebrows in face recognition.Perception, 32, 285-293.

研究23 观察你的情绪了吗

Ross, P.(2003, August 11).Mind readers.Scientific American, 289(3), 74-77.

这篇研究报告与本书中的其他研究相比不尽相同，但将之收录进来似乎很合适。如你所知，本书涉及那些意义非凡、改变心理学的传统研究。本节的研究也不例外，然而，或许你已注意到，该研究完成于近期。目前，有一类研究正以一种极具说服力的方式改变着心理学，这项研究正是它们的代表。这类研究涉及新兴的、具有高度技术性的方法，该方法可以通过无创的方式真正窥视你的大脑，“看见”大脑中正发生什么，以及大脑如何执行诸如情绪、思考和感受等功能。好吧，这也许并非货真价实的“读心术”，但如今心理学与神经科学的研究技术，其令人震惊的程度并不亚于“读心术”。

大多数人都听说过核磁共振成像（MRI），很多人也许还接受过扫描。与X光和CT扫描相比，MRI是一项重大进步，它不发出放射线，因此对人体无害，并且它能比CT扫描更清楚地显示软组织异常（它的图像通常是三维的，这对大脑结构的可视化尤其关键）。因此，如果你膝盖受伤，CT扫描未见任何明显损伤时，下一步就极有可能进行MRI扫描，以便确切查看是否发生了软骨或其他组织损伤，从而对症治疗。

从20世纪70年代开始，MRI就已经被运用于医疗领域，然而在过去40年间，这项技术发生了极大的演变。开发MRI成像技术的先驱是伊利诺伊大学的保罗·劳特布尔（Paul Lauterbur）和英国诺丁汉大学的彼得·曼斯菲尔德（Peter Mansfield）（Lauterbur，1973；Mansfield，2007）。鉴于在这项技术上的杰出工作，他们共同获得了2003年的诺贝尔医学奖。



图6—1 fMRI扫描舱（Juan Manuel Silva/Glow Images，Inc）

在此展示菲利普·罗斯（Philip Ross）的一篇早期研究报告，该报告探索了如何将MRI用于医学用途以外，比如，探测涉及人类思维、情绪、动机和行为的大脑结构。具体来说，罗斯进行了一项研究，探索MRI是否能判断人在说谎——它是否有可能成为一台高科技测谎仪！

运用MRI“读取”大脑的基础是，在扫描过程中的某一时刻，它能显示大脑的哪些部分正在运转。MRI的这一探测大脑功能运转而非仅作诊断工具的用途，被称为fMRI（功能性核磁共振成像，如图6—1所示）。换言之，当某人的大脑被MRI扫描时，研究者可以让他进行诸如回答问题、观看图片、阅读文段等活动。随后，研究者能准确地了解大脑中的哪些部分参与了上述活动。这是如何做到的呢？其实，MRI是通过探测大脑激活区周围血流中的氧含量和葡萄糖消耗量的变化，来了解大脑是如何工作的。研究者通过扫描仪的监视器，观察大脑中某些部分的活动

变化，他们能大致看出大脑中哪些部分在完成任务时被激活了。从某种程度上说，这些脑区被“点亮”了。

理论假设

试着回答一下，当人讲真话或撒谎时（人人都会撒谎；有些人撒谎比别人多），真话和谎言是否“来自”大脑中不同的部分？个人感受显然不能帮助我们回答这一问题。你的大脑并未感到任何不同。但如果我们能观察大脑在讲真话和撒谎时的运转情况，看看针对每一项活动，哪些区域激活了，我们就能回答这个问题。

fMRI能做到这点，并且它比多导生理记录仪（即测谎仪）更胜一筹，因为多导生理记录仪主要测量除大脑以外的且由焦虑引起的身体反应，而人们产生焦虑可能出于各种各样的原因，不仅是由于撒谎。因此，科学上并不认为多导生理记录仪是有效的，法庭上也不允许其作为呈堂证据。fMRI脑成像并不从研究焦虑出发，而是揭示思考过程。

这篇文章的基本观点是，在扫描仪中对大脑进行观察，如果被试说了真话，其大脑的一个部分将“点亮”：这个部分负责真实内容的陈述。然而，如果此人撒谎，大脑中的两个部分将显现出来：一部分制造谎言，另一部分抑制被谎言所掩盖的真话。换句话说，谎言是掩盖真话的，否则它就无法存在，这两个过程处于大脑中的不同部分（见Langleben et al.，2002）。使用fMRI应当可以探测这些脑区。

方法

文中描述的这项关于说真话和撒谎的研究，是在被试和研究者都清楚哪些是真话、哪些是谎话的情况下进行的。换言之，这并非用fMRI来检查某个人——如被控犯罪或被质疑失德和行为不端者——是否撒谎的测试。这其实是一项用以检验理论的探索性研究（Langleben et al.，2002）。

方法非常简单。要求被试从三只信封中选择一个信封，每只信封内都装有梅花5扑克牌和20美元钞票（被试并不知道信封中的物品都是一样的）。随后，要求被试记住牌，并将牌和钱放回信封，将信封藏在他们的衣袋里。他们被告知，有一台“电脑”能分析人的想法，如果被试能在电脑前成功隐瞒自己所持有的牌，就能获得信封里的20美元（实际上，除了fMRI设备以外，并未使用所谓的电脑）。另一些被试拿到了红心2，他们被指示要说真话，不要隐瞒他们持有的牌。

随后，每名被试都被置于fMRI扫描仪中，可使用是/否按键。他们

被问道："你有这张牌吗？"被试并不知道研究者清楚他们持有哪一张牌（这一点无关紧要，因为被试"欺骗"研究者不是本研究的重点）。

结果

fMRI扫描的发现是清楚的。如你所见，在图6—2中，讲真话的被试只运用了一个脑区（前扣带回）。然而，在撒谎条件下，两个区域（前扣带回和背侧前额叶皮层）活动了起来。

研究意义

这些发现之所以意义重大，主要是基于两个原因。

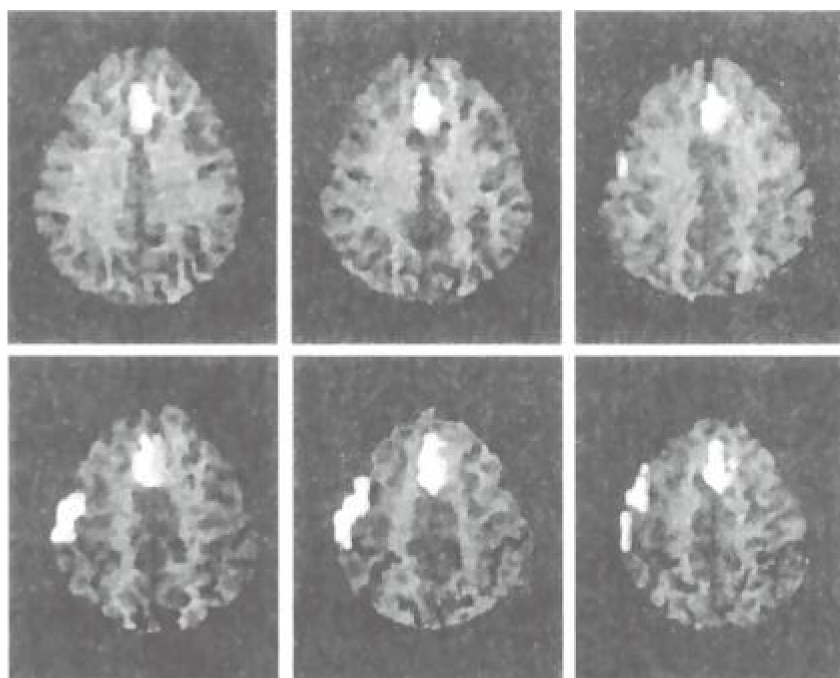


图6—2 讲真话和撒谎被试的fMRI扫描结果

注：上图：讲真话（红心2）被试的脑扫描；下图：撒谎被试的脑扫描（梅花5）。

资料来源：见Langleben et al., 2002。

第一，它们证实了MRI扫描可能具有医学诊断之外的重要用途（它确实如此，很快会讨论到），并且也许能成为一架"傻瓜式"测谎仪，其

结果最终可能被法庭承认，正如DNA证据在证明或证伪某些犯罪方面，已经变为一个可被接受的方法。

第二，这项研究扩宽了我们对于人类大脑那不可思议的特异性的认识。如果你直视一个人的大脑，它看起来不外乎是一团又软又稠的细胞和组织。但事实上，你的大脑高度特异化，它为具体的任务备有具体的区域。你的大脑拥有为数众多的特异性脑区，它们被“分派”了从面孔识别到说话、计划和问题解决等具体工作。在这项研究中，你看见了一个致力于讲真话的脑区和另一个致力于撒谎的脑区。

近期应用

自21世纪初这篇文章发表以来，心理学在“读脑”方面有何进展？我们如今能把fMRI用作测谎仪吗？fMRI还能让我们做别的什么事？就第二个问题而言，答案似乎是“可能”。一项研究发现，fMRI测谎比传统的生理多导测试显然更加可靠。不过，反对者指出，法律允许窥视大脑可能引起严肃的伦理学问题（Simpson，2008）。

另一项值得注意的重要研究，运用了和在此讨论过的玩牌研究相同的方法，但使用了数组真实的人类面孔（Bhatt等，2008）。向被试呈现数组面孔照片，随后在fMRI的扫描下，从他们刚才所看过的面孔照片中选出一些来呈现，要求他们或如实报告此前曾见过这些照片，或撒谎。结果与扑克牌实验基本相同。两个脑区在他们撒谎时被激活，但仅一个脑区在讲真话时被激活。

fMRI在“测谎仪”之外用途的一个例子和心因性疾病有关。在神经心理医学中，fMRI已经成为强大助力，它为精神分裂症患者脑中可能导致理性思维特征性缺失的脑区进行定位（Libby & Ragland, 2011）。这类定位信息或许能帮助研究者开发新疗法，以期应对精神分裂症这一最具毁灭性的疾病。神经心理学通过运用fMRI取得的另一项重大进步在于对自闭症谱系障碍个体脑功能认识的提升（Philipa et al., 2012）。这类障碍通常被清楚地标记为社交缺陷，即不能或难以与他人恰当地建立关系。一项针对自闭症患者的fMRI研究显示，负责社会化的脑区有异常活动，这一结果提示自闭症患者可能并非对社会化无能，而是对社交回避有一种偏好。该研究还揭示了当常规任务发生变化时，大脑如何进行反应调节，这也许能解释为什么与自闭症有关障碍患者在他们所遵循的习惯被改变后，会显得心烦意乱。

2012年，一项也许是最令人惊讶的fMRI结果出现了，它来自加州大学伯克利分校的加兰特fMRI实验室（Gallant fMRI Laboratory at U.C.Berkeley, California）。在fMRI扫描室中，给被试呈现电影片段。运用新的电脑软件，研究者得以重建扫描过程中被试所观看过画面的图片，图片虽然模糊却具有极为典型的特征。如图6—3所示，对照电影截

图，你会发现这项技术和真正的读心术有多么接近。



图6—3 电影截图及其fMRI扫描下的重制图

注：左上和左下：AF archive/Alamy; 右上和右下：Vadym Drobot/Shutterstock。

结论

fMRI的效用——比如帮助我们明白大脑是如何工作的：它是否与测谎、心因性脑功能失调和脑损伤相联系，或者它是否能揭示个体或群体（例如男人和女人）的大脑是如何以不同方式进行连接的——在很大程度上尚未实现。但无论如何，在诸如MRI这类无创扫描技术之前，我们直接研究大脑的方法还局限于外科手术和尸体解剖。现在，神经学家可以运用fMRI和其他无创工具，比以往更加有效地研究、诊断、治愈脑部相关疾病，甚至还可能记录你的思想。这项技术究竟会被用于善意还是“邪恶”的用途，将取决于操控者的伦理观和价值观。

Bhatt, S., Mbwana, J., Adeyemo, A., Sawyer, A., Hailu, A.& VanMeter, J.(2009).Lying about facial recognition: An fMRI study.Brain and Cognition, 69, 382-390.

Langleben, D, Schroeder, L., Maldjian, J., Gur, R., McDonald,

S.Ragland, J., O’Brien, C., & Childress, R.(2002).Brain activity during stimulated deception: An event-related functional magnetic resonance study.NeuroImage，15, 727-732.

Lauterbur, P.(1973).Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance.Nature 242，190-191.

Libby, L.& Ragland, J.(2011).fMRI as a measure of cognition related brain circuitry in schizophrenia.Current Topics in Behavioral Neurosciences.Retreived from www.biomedsearch.com/nih/fMRI-as-Measure-Cognition-Related/22105156.html.

Mansfield, P.(1973).A personal view of my involvement in the development of NMR and the conception and development of MRI.Nature, 242，190-191.

Philipa, R., Dauvemanna, M., Whalleya, H., Baynhama, K.Lawrie, S., & Stanfield, A.(2012).A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders.Neuroscience & Biobehavioral Reviews, 36, 901-942.

Simpson, J.(2008).Functional MRI lie detection: Too good to be true? Journal of the American Academy of Psychiatry and Law, 36, 491-498.

研究24 认知失调

Festinger, L., & Carlsmith, J.M.(1959).Cognitive consequences of forced compliance.Journal of Abnormal and Social Psychology, 58, 203-210.

你是否有过必须说或做与自己意愿相反的事情的尴尬境遇？很有可能！每个人都会碰到这种情况。当你不得不这么做时，你真正的态度或观点会发生什么变化？还是没有什么变化？也许它的确没有发生变化。然而，研究发现，当你的行为和你的态度发生矛盾时，你的态度将有所改变，以与行为保持一致。例如，如果因实验要求而强迫一个人发表讲话，并支持一种与他原来的观点相悖的意见，那么他的真实态度也将渐渐转向他在讲话中所支持的观点。

20世纪50年代初，许多研究把这种观点的改变归结为：（a）该次演讲的心理演练复述（mentally rehearsing）；（b）从相反立场寻找论据的结果。这些早期的理论认为，在完成上述心理任务的过程中，被试让他们自己相信了他们在讲话中所持的观点和立场。为了做进一步的推论，对做出与自身所持观点相悖却令人信服的讲演的被试，研究者给予了金钱奖励。人们预测被试所得到的奖励越多，他自己的观点就越容易改变。这似乎很符合逻辑，不是吗？然而，这又一次证明常识不能很好地预测人类行为。这些研究发现，事实与人们预期的相反，即给被试的奖励越多，他们的态度改变得越少。从当时盛行的学习理论（操作性条件反射，强化理论等等）角度看，研究者很难对上述发现作出合理的解释。

不久之后，斯坦福大学的心理学家莱昂·费斯廷格（Leon Festinger）提出了颇具影响力且至今闻名于世的认知失调理论（cognitive dissonance），用该理论便可解释上述看似矛盾的研究发现。“认知”一词指的是一种心理过程，如思想、观念、知识单元、态度或信念等；“失调”的意思就是不协调（out of tune）。基于此，费斯廷格主张，同时经历两种或多种彼此间不协调的心理过程，就是认知失调。一旦出现了这种情况，它将导致不同程度的不适感和紧张感，其程度取决于该认知失调对你的重要程度。进而，这种不适感会促使你做出某种改变，以使它减弱。既然你不能改变你的行为（因为你已经这么做了，或是因为环境的压力太大），你只得改变你的态度。

费斯廷格创立这一理论的灵感来源于1934年印度发生的一场大地震。在这场地震后，可怕的谣言传遍整个印度。其内容是在此次地震灾区以外的地区将会有范围更广、震级更大的地震发生。这些谣传没有任何科学证据，但令费斯廷格不解的是，人们为什么会散布这种灾难性的、令人焦虑的谣传呢？不久，他恍然大悟：谣传不是用来增加焦虑

的，而是用来为焦虑辩护的。也就是说，即便他们住在灾区外，这些人还是感到非常害怕。这就产生了认知失调，即虽然缺乏科学证据，但他们的这种害怕却无法调适。因此，他们散布这种谣传以证明他们的害怕是事出有因的，进而也就降低了他们的认知失调。他们使自己对世界的看法与自身的感受和行动相一致。

理论假设

费斯廷格的理论观点是：通常情况下，在我们的社会里，个人所公开表达的看法与其私下里的观点或信仰是相一致的。因此，如果你相信X，但你却公开主张非X，那么你将会体验到这种由认知失调引起的不适感。不过，如果你识别到自己之所以公开主张非X是由于某种不得已的原因，如压力、奖励或者惩罚，那么你的认知失调将大大降低或者消失。因此，关键在于你越将不一致的行为归因于自己的选择，你的认知失调就会越严重。

对你而言，降低这种令人不快的认知失调的方法之一是改变自己原有的观点，使之与自己的行为（即某种新主张）相一致。费斯廷格认为，当个体的认知失调很严重时，其态度和观点的改变幅度将是最大的。设想一下，假如某人给你一大笔钱，让你公开支持某种你原本反对的观点，而且你也同意这么做。再假设另外一个人向你提出同样的请求，但只给你一小笔钱，虽然这看起来完全不值得一做，但不管怎样，你还是答应了。那么，在哪种情况下，你的认知失调会更大呢？从逻辑上讲，在钱较少的情况下，你将经历更大的认知失调，这是因为你没有充足的理由为你那违背自己意愿的言行辩护。因此，根据费斯廷格的理论，在钱较少的情况下，你真正的观点将会有更大的改变。现在，让我们来看看费斯廷格和他的助手詹姆斯·卡尔史密斯（James Carlsmith）是如何设计实验来验证自己的认知失调理论的。

方法

假设你是一名大学生，正在学习一门心理学导论课。该课程的要求之一是你必须在这一学期中以被试身份参加三个小时的心理学实验。你查看了布告栏，那里贴着许多心理学教授和研究生正在进行的心理学研究，随后你在一个耗时两个小时的一项行为测量（measures of performance）的研究上签了名。这一研究由费斯廷格和卡尔史密斯主持，如许多心理学实验一样，该研究的真正目的对被试保密，因为如果被试知道了研究的真正目的，其反应将会出现偏差，进而使研究结果失效。实际的被试由71名正在学习心理学的低年级男生组成。

你在规定的时间来到实验室（这里的实验室就是一间有些椅子的空

房间）。工作人员告诉你这个实验需要花费一个多小时的时间，因此必须做两个小时的打算。既然还有一些剩余时间，实验者会告诉你，一些心理系的人正在就作为被试的感受这一话题，访谈曾做过被试的大学生，而且让你参加完这项实验后也去接受访谈。随后实验者便给你布置了第一项任务。

你的面前放着一个装有12个线轴的托盘，主试要求你从托盘里拿出线轴放在桌上，然后再将线轴放回原处，并且如此循环往复。你要用一只手以你自己的速度完成任务。与此同时，实验者拿着秒表在旁观看并做记录。就这样，你必须反反复复做30分钟。然后，主试拿走托盘，又给了你一块钉有48个方栓的板，你的任务是按顺时针方向将每个方栓旋转90度，循环往复再做30分钟！如果你觉得这种活动太过枯燥，那这正是该研究所欲达到的效果。按费斯廷格和卡史密斯的话说："该实验阶段的目的就是要使所有被试产生一种一致的消极体验。"（P.205）毫无疑问，你会承认这一目的已经达到了。完成了这些任务后，真正的实验才刚刚开始。

主试将被试随机分配到三种实验条件下。在控制条件下的被试完成上述任务后即被带进另一个房间，就自己对刚刚完成的实验任务的反应接受访谈。其余的被试接着进行下一步实验程序。完成了上述任务后，实验者与实验组的被试进行谈话，好像是要向他们解释该研究的目的。实验者对每位被试都说，他是A组的被试，他们这组在实验前没有得到与实验有关的任何信息。还有一个B组，对B组的被试，实验者会在实验前向他们传达这样的信息——这些实验任务是非常有趣的，而这一信息是由另一名大学生来传递的，他伪装成已完成了实验任务的被试。请记住，上述一切都是刻意布置的，为的是确保随后实验的关键部分真实可信。换句话说，这些解释仅仅是一种托辞。

然后实验者离开房间几分钟。而当他返回时，表情略显迟疑和困惑，并继续对被试说，那名向B组被试传递信息的大学生打电话来说他病了，但现在还有一名B组被试在外面等待，而他们这会儿又很难再找一个人填补这一空缺。然后，实验者很有礼貌地问被试是否愿意顶替这名大学生去向正在等候的B组被试传递信息。

实验者会给那些愿意提供这一帮助的一些被试，每人1美元作为酬谢，而给另一些愿意提供帮助的被试，每人20美元。被试同意参加下一步实验后，实验者便交给他一张纸，上面标有"B组"几个字，并写着这样一段话："这项实验任务非常有趣，我从中获得了快乐，我很高兴，这是激动人心的，令人兴奋的。"然后实验者付给被试1美元或20美元，并将他带入一间房间和新来的"被试"在一起单独呆上两分钟。两分钟后，实验者回到这间房间，对被试说了些表示感谢的话，然后把被试带入访谈室，让他们谈谈对实验任务的真正看法。这一访谈与先前对控制组进行的访谈完全一样。

整个实验程序是否略显复杂？其实一点也不。说得简洁明了一些就

是，共有三组被试，每组各有20人：一组被试因其就实验任务向后来的“被试”撒谎而得到1美元，另一组被试因此而得到20美元；控制组被试没有撒谎。

结果

研究结果以每名被试在最后访谈阶段所表达的对实验任务的真实感受作为测量指标。要求他们对该实验做如下评定：

1.这些实验任务是有趣且令人愉快的吗？以11点量表的形式进行评定：–5代表非常枯燥以及使人厌烦；＋5代表非常有趣且令人愉快；0代表这些任务是中性的，即不枯燥也不有趣。

2.你对自己完成这些实验任务的能力了解了多少？以11点量表的形式进行评定：0代表毫无收获；10代表收获非常大。

3.你相信该实验及你所完成的实验任务是在测量某些重要的东西吗？以11点量表的形式进行评定：0代表没有什么科学价值；10代表科学价值非常大。

4.你是否还愿意参加另外一项与之类似的实验？以11点量表的形式进行评定：–5代表肯定不想参加；＋5代表非常想参加；0代表中立的看法。

表6—6中列出了所有被试对每个问题所作回答的平均值。问题1和问题4是用以说明费斯廷格的认知失调理论的，其结果表明组间差异非常显著。相对于那些得到20美元及没有撒谎的被试而言，得到1美元的被试更喜欢这些实验任务；这与此前该领域的一些研究结论相反，而且也和我们大多数人用常识对结果所进行的预料相反。这一研究发现既反映在对第一题（直接问题）的回答上，也反映在获得1美元的被试身上，他们表示更愿意参加另一项类似的实验上（即对问题4的回答）。

表6—6 每种实验条件下的被试对访谈问题回答的平均评定值

问题	控制组	1美元组	20美元组
1. 任务是否有趣（–5到+5）*	0.45	+1.35	0.05
2. 有多大收获（0到10）	3.08	2.80	3.15
3. 科学价值如何（0到10）	5.60	6.45	5.18
4. 是否愿意再参加类似的实验（–5到+5）*	0.62	+1.20	0.25

*与费斯廷格和卡史密斯的研究假设有关的问题（引自P.207）

讨论

用费斯廷格的话说，认知失调理论主张如下。

1.如果一个人被别人劝说做了某种和他原来的观点相反的事或说了某些与他的真实观点的话，他将会产生一种改变原有观点的倾向以使之与刚刚做过的事或说过的话相一致。

2.引起撒谎行为的压力越大，则改变原有观点的倾向将越弱（P.209—210）。

费斯廷格和卡尔史密斯的研究发现显然支持了这一理论。费斯廷格对此的解释是：当人们从事某种与态度不符的行为（撒谎）时，如果他们有这样做的充分理由（20美元），他们将体验到较少的认知失调，因此他们改变观点的动机并不强烈。相反，如果他们的理由并不充分（1美元），他们的认知失调将比较严重。因此，为了减少认知失调所引起的不适感，他们对原有的观点将作出较大的调整。

该理论可以用图解的形式表述如下：

与态度不符的行为→解释该行为的理由充足→轻微认知失调→态度改变较小

与态度不符的行为→解释该行为的理由不充足→严重认知失调→态度改变较大

质疑与批评

费斯廷格预料到，许多早期理论会受到这一新观点的威胁，它们的支持者会试图批评这一新的研究结论，并提出他们自己对这一研究结果的解释（诸如我们在本章开头所提到的心理演练及寻找论据等）。为了反击这些批评意见，该研究对实验组被试撒谎的实验片段进行了录音，并且由两个并不知情的评定者各自单独进行评定。统计结果显示，两组被试（1美元组和20美元组）在撒谎的内容或说服力方面不存在显著差异。因此，唯一剩下的对该研究结果的明确解释，便是费斯廷格的认知失调。

自从费斯廷格和卡尔史密斯提出认知失调理论以来，许多其他研究者对该理论进行了修正，但并不否定该理论。库珀和法齐奥对各种修改意见进行了总结（Cooper & Fazio，1984），他们还提出了由认知失调所引起的态度改变必须经过4个必要的步骤。

第一步，这种与态度不符的行为必须导致你不乐于接受的消极结

果。在费斯廷格和卡尔史密斯的实验中，被试不得不向“新来的被试”撒谎，以说服他们参加一个令人厌烦的实验，这就导致了所需的消极结果。这也就是为什么当你恭维某人的穿着，而实际上你根本无法忍受他的穿着时，你对他的穿着的态度没有改变的原因。

第二步，个人必须对消极结果承担责任。这通常涉及是否有选择权的问题。如果与态度不符的行为是你自己的选择，且导致了消极结果，那么你将体验到认知失调。但如果有人强迫你这样做，你将不会感到自己对此有责任，因而也就不会产生认知失调。尽管费斯廷格和卡尔史密斯在论文的标题中使用了“强迫服从”（forced compliance）一词，但实际上被试仍相信其行为是自愿的。

第三步，研究表明，生理唤起（physiological arousal）也是认知失调过程中必不可少的组成部分。费斯廷格认为，失调是一种促使我们改变态度的令人不适的紧张状态。的确，有研究已经证明，当被试通过与态度不符的方式自主行事时，他们体验到了生理上的唤起。尽管费斯廷格和卡尔史密斯没有测量被试的生理唤起，但可以肯定地说，生理唤醒是存在的。

第四步，个体必须意识到自己所体验到的唤醒是由自己所做的与态度不符的行为引起的。在费斯廷格和卡尔史密斯的研究中，被试可以轻易体验到的不适感可以明确地归结为他们向新来的被试撒谎所引起的。

费斯廷格和卡尔史密斯的认知失调理论已经成为被人广泛接受且经得起检验的心理学研究成果。大多数心理学家都认为，有两个基本过程足可以使我们改变自己原有的观点和态度：其一是劝说，即其他人不遗余力地说服你改变自己的观点；其二就是认知失调。

近期应用

费斯汀格和卡尔史密斯的研究理论和结果不断得到社会科学研究领域的支持和肯定。一项有趣的研究发现，仅仅是由于观察到你喜欢或是尊敬的人做出与你的态度不一致的行为，你就会体验到认知失调，并改变你自己的态度，完全不需要亲身参与（Norton et al., 2003）。作者把这个过程称为替代性失调（vicarious dissonance）。在这项研究中，大学生在一个有争议的问题上听到一些与他们态度相反的演讲（关于增加大学学费的问题）。其中一些大学生，听到的关于支持增加大学学费的言论是来自于他们自己学校（“内群体”），另外一些人听到的言论是来自于其他大学中的学生（“外群体”）。当内群体成员提出此言论时，被试感到认知失调，而在增加学费的问题上的相反态度有所减弱。另一个研究结果更有力地显示出替代性失调，研究者发现被试根本不必亲耳听到某些言论，只简单地了解到内群体成员同意发表支持学费增加的演讲，就会产生失调，进而引起态度改变。

颇具吸引力的另一项研究在另一个完全不同的领域使用了认知失调理论，它解释了为什么吸烟者在明知吸烟有害健康（几乎每个人都知道）的情况下仍不能戒烟的原因（Peretti-Watel等人，2007）。如果你吸烟，并且知道吸烟的危害，但却感觉无法戒掉，那你一定有过认知失调的经历。因为这不是一个令人愉快的状态，因此，你会制定一个减轻你不适感的策略。在2007年的研究中，研究人员发现吸烟者经常沿着这样的思路来表达一种“自我豁免”的信念，比如“吸烟有害健康，但我除外，因为我不经常吸烟”或者“我吸烟是为了保护自己远离疾病”。研究人员建议：“未来的控烟信息上，应该有专门针对那些降低吸烟者的认知失调和抑制他们戒烟意愿的自我免疫信念的干预措施。”（P.307）。

加州大学圣克鲁兹分校的心理学家艾略特·阿伦森（Elliot Aronson）完成了一项以费斯廷格的认知失调理论为基础的重要研究，该研究的目的是改变学生的危险性行为（Shea，1997）。研究者要求热衷于性行为的学生就“使用避孕套能减少感染艾滋病的概率”为主题，做一段宣传录像。做完录像后，其中一半的学生被分为多个小组，研究者鼓励他们就“为什么大学生不必使用避孕套”的问题展开讨论，并透露他们自己不使用避孕套的经历。换言之，这些被试不得不承认他们并不总是坚持刚刚自己在录像中所宣扬的立场。这样，他们便不得不面对自己的言行不一。参与录像制作的另一半学生没有参与后面的讨论。当研究者给所有学生一个购买避孕套的机会时，与未参与讨论的学生相比，自相矛盾组的学生购买避孕套的人数更多。更重要的是，3个月后，当被试就他们的性生活接受访谈时，92%的自相矛盾组的学生声称他们在性交时一直使用避孕套；而在未参与讨论组，只有55%的学生声称自己一直使用避孕套。这是一个认知失调起作用的典型例子。

结论

当你不得不面对自己的言行不一的情况时，通常都会体验到认知失调，这将促使你改变自己的行为或观念来使自己的言行保持协调一致。阿伦森坚信：认知失调在改变现实行为方面具有重要作用。他解释说：“我们大多数人经常会有言行不一的时候，这是因为我们会对这样的行为熟视无睹。但是如果有人站出来，而且迫使你正视它，你可能就不会再对其置之不理了。”（Shea，1997，P.A15）

Cooper, J., & Fazio, R.(1984).A new look at dissonance theory.In L.Berkowitz (Ed.), Advances in experimental social psychology.New York: Academic Press.

Norton, M.I., Monin, B., Cooper, J., & Hogg, M.A.(2003).Vicarious dissonance: Attitude change from the inconsistency of others.Journal of Personality and Social Psychology, 85, 47-62.

Peretti-Watel, P., Halfen, S., & Gremy, I.(2007).Risk denial about smoking hazards and readiness to quit among French smokers: An exploratory study.Addictive Behaviors, 32, 377-383.

Shea, C.(1997).A University of California psychologist investigates new approaches to changing human behavior.Chronicle of Higher Education, 43(41), A15.

第七章 人格

当你问自己“我是谁”的时候，你恰好问了一个人格心理学家提出的基本问题。人格心理学家们力图揭示那些使每个人独一无二的人格特征，并确定这些人格特征的起源。当行为科学家谈到人格的时候，他们通常指的是那些相对稳定的、具有跨情境和跨时间一致性的品质。也就是说，你是一个什么样的人，不会每天、每周发生改变，而且常常是每年甚至数十年都不会改变。相反，你的某些基本的人格特征是恒定且可预测的。在心理学历史上，心理学家已经提出几百种人格理论。其中绝大多数理论模型都引发了很多讨论和争论，人们并不清楚这些模型是否真正测量了个体之间的重要差异。当然，经反复验证，有一些因素能够可靠地预测某些特定行为。这些因素就是本章所要讨论的重点。

第一篇文章讨论的是朱立安·罗特（Julian Rotter）的著名研究，它主要关注的是人们对生活中控制点的看法。一些人相信他们的生活是被外部因素所左右的，比如命运或运气，而另一些人则认为他们的生活是被内部因素所操控的，即他们是自己生活的主宰。一个人关于外部或内部控制的信念品质已被证明是反映个人人格特征的一种具有一致性和重要性的因素。接下来的文章是关于桑德拉·贝姆（Sandra Bem）在20世纪70年代的研究，该研究颠覆了我们对于性别这一人格同一性的基本成分的看法。第三篇文章所介绍的研究影响广泛：该研究率先确定了所谓的“A型”与“B型”人格，以及这两种类型的人有什么根本差别。有很多原因能说明这些差别的重要性，其中一个原因就是“A型”个体可能更容易患冠心病。你还会读到一篇在心理学各领域颇具影响力的研究。该研究提醒我们必须把人类行为放在某种文化背景下加以考虑。这篇研究探讨了哈里·川迪斯（Harry Triandis）的工作，他在过去三十年间，经仔细探索构建了令人信服的理论，即认为我们可以将大多数人类社会划分为集体主义文化和个人主义文化两种。这种看似单一（尽管肯定不简单）的维度可以在很大程度上解释了文化背景对于个体的人格特质所造成的深远影响。

研究25 你能主宰自己的命运吗

Rotter, J.B.(1966).Generalized expectancies for internal versus external control of reinforcement.Psychological Monographs, 80，1-28.

你的行为后果是由你本人控制，还是由外在力量所左右？请你思考一下下面这些问题：当某种好事发生在你身上时，你认为这是实至名归的呢还是运气所致？而当某种不好的事情发生时，你会觉得这是由于自己的行为所导致的呢还是把它归结于命运？这一问题可以用更正式的心理学术语表述如下：你相信在你的行为选择及其后果之间存在某种因果关系吗？

作为心理学历史上最有影响力的行为主义者之一，朱利安·罗特认为个体在某些事件原因的归因方面存在很大的差异。当人们把自己的行为后果归结为运气、命运或是他人力量的影响时，表明他们持有罗特所说的外控点（external locus of control）的信念；与之相反，如果人们把行为后果归因于自己的选择和人格，则他们就具有一种内控点（internal locus of control）的信念。在他这篇于1966年完成并被频繁引用的论文中，罗特解释说："一个人看待事件的内外倾向是决定我们是怎样的人的基本因素，而且这可以用社会学习理论加以解释。"

在社会学习理论看来，一个人在从婴儿成长为儿童的这段时间内，会受到某种形式的奖赏或强化而习得一些特定情境中的行为。这种强化增加了儿童的预期，认为某种特定行为将产生某些他们想要的奖赏。这种预期形成之后，取消强化将导致预期渐渐消退。因此，有时人们认为强化与行为是相关联的，有时却不是这样（参见研究11中斯金纳对关联性问题的讨论）。随着儿童的成长，一些儿童经常感受到其行为与强化的直接联系，而对其他儿童而言，结果似乎并非由其自身的行为所导致。罗特称，个体所有的具体学习经历使其对强化是内控还是外控而形成了一种总体预期。

罗特写道："随着在文化上将情境类分为社会决定型情境和技能决定型情境，这些总体预期将导致行为的特征差异，并进而在特定条件下，演变为个体差异。"（P.2）换言之，你已经习惯于对自己的行为后果做内控或外控的解释，这将影响到将来你在几乎所有情境中的行为。罗特坚信，不论是外控还是内控，你对控制点的解释始终是你个性特征的重要组成部分，也是你人格的一部分。

现在，让我们回过头来再看本文开头所提出的问题。你认为自己是内控的还是外控的？罗特试图研究人们在这一维度上的差异，当然他不只是简单地对被试进行询问，而是编制了一套用以测量个体的控制点的

测验。一旦测量出人们的这种特征，就能进而研究该人格特征对个体行为的影响了。

理论假设

罗特打算在他的研究中证明两个要点：首先，他预言可以研发一个测验，可靠地测量出个体在生活中归因的内控或外控倾向的程度；其次，他假设人们在同一情境中对强化原因的解释将呈现出稳定的个体差异。罗特计划通过比较“内控者”和“外控者”在各种情境中的行为来验证他的假设。

方法

罗特设计了一个包括一系列成对陈述句的量表。每对句子中包括一句反映内控点的陈述句和一句反映外控点的陈述句。给被试的指导语是：“从每一对陈述句中挑选出一个句子，它所描述的情况必须是你认为自己在身临其境时更有可能发生的情况，即你一定要选择那个对你来说更真实的陈述句，而不要选择你认为应当选择或希望它成为事实的陈述句。”这是一种对个人信念的测量，回答显然没有对或错之分（P.26）。该测验采用迫选的回答方式，即迫使被试在每对陈述句中选择其中一句，既不能不选，也不能两个都选。

罗特的这一测量工具几经修订和改进。早期的量表包括60对陈述句，但经过各种信度和效度的检验后，量表的条目被精简到23项。为了掩盖该测验的真正目的，后又加入了6个补充项目。因为如果被试能够猜到测验想要测量的内容，他们可能在某种程度上改变自己的回答以求“表现得更好”，所以在心理学测验中常常用到类似的补充项目。

罗特称他的测验为I-E量表（“I”指的是Internal，内控；“E”指的是External，外控），这也是如今该测验广为人知的名称。表7—1列出了一些取自I-E量表的样题，以及补充项目的样题。如果你仔细看一下这些项目，就会相当清楚哪些陈述句反映内控倾向，哪些陈述句反映外控倾向。罗特声称他的测验所测量的是一个人的人格特征中内控或外控的程度。

罗特接下来要做的，也是最为重要的一步就是证明他可以用内—外控人格特征准确预测人们在特定情境中的行为。为此，罗特报告了几项研究（由他本人和其他人共同完成）来考查被试在I-E量表的得分与他们在各种生活事件中的个体行为之间的关系。这些研究表明，I-E得分与人们在很多不同情境中的行为都存在显著相关，例如赌博、政治活动、劝说、吸烟、成就动机以及从众倾向等。

表7—1 罗特I-E量表的样题和填充项目

项目#	陈述句
2a	人们生活中很多不愉快的事件都有一部分原因在于运气不好。
2b	人们的不幸是由他们所犯的错误导致的。
11a	成功靠的是努力工作，运气只起一点点作用或者根本不起作用。
11b	找到一份好工作主要靠的是在恰当的时间出现在恰当的地点。
18a	大多数人没有意识到他们的生活在某种程度上是被偶发事件所左右的。
18b	世上根本就没有“运气”这回事。
23a	有时我不能理解老师是怎样给我打分的。
23b	在我的努力程度和学习成绩之间有直接的关系。
	补充项目
1a	孩子们之所以闯祸，是因为父母对他们的惩罚太多。
1b	今天的孩子所面临的困难是他们的父母对他们过于随和。
14a	世上有一无是处之人。
14b	每个人都有一些优点。

（引自P.13 14）

结果

下面我们对罗特在该领域研究中的早期发现做一简要总结（完整讨论以及参考文献参见原文第19-24页的）。

赌博

罗特研究了赌博行为与控制点的关系。研究发现，经I-E量表测得的内控型个体倾向于对“有把握的事情”下赌注，且相对于小概率事件而言，他们更喜欢中等概率；相反，外控型个体更愿意在风险较大的赌博中下更多的注。此外，外控型个体在下注时更倾向于采用被人们称为“赌徒谬论”的异常方式（例如，即使实际发生的概率不变，但还是认为某个很久没有出现的数字“应该”会出现，从而在这个数字上下更多的赌注）。

劝说

罗特引用了一项有趣的研究，该研究用I-E量表选取了两组学生，一组为高度内控者，另一组为高度外控者。两组被试对校园里的男生联谊会和女生联谊会大体上持相似的态度。主试要求两组被试劝说其他的学生改变对上述组织的看法。结果发现，内控组被试比外控组被试做得更为成功。相反地，另外一些研究发现内控型被试更不愿在态度方面受他人的影响。

吸烟

内控倾向似乎与自律存在相关。罗特提及的两项研究均发现：（a）吸烟者比不吸烟者更倾向于外控；（b）那些在1966年烟盒上出现有关吸烟危害健康的警示之后，成功戒烟的个体更多地倾向于内控，尽管不论是内控型个体，还是外控型个体，他们都相信这种警示是正确的。

成就动机

如果你相信自己的行动决定着自己的成就，那么从逻辑上讲，你应该比那些认为成功更多取决于命运的人有更强的成就动机。罗特对1000名高中生所做的研究表明，I-E量表中的内控得分与成就动机存在正相关。这些成就动机的指标包括上大学的计划、用于做家庭作业的时间、父母对子女的在校表现的感兴趣程度，等等。这些与成就动机有关的因素更可能从具有内控倾向的学生身上发现。

从众

罗特引用所罗门·阿希（Solomon Asch）所做的从众实验来证明自己的观点。在该实验中，主试让被试处于一个大多数人都认可同一个不正确答案的情况下，以判断被试的从众行为（参见研究38中阿希的从众实验）。主试允许被试对他们判断的正确性下赌注（钱由实验者提供）。在这样的条件下，与外控倾向的被试相比，那些具有内控倾向的被试更少服从大多数人的意见，而且当他们的判断与大多数人相反时，他们下的赌注更多。

讨论

罗特在论文的讨论部分提出了人们在内外控维度上存在个体差异的可能原因。他引用了多项研究指出，内、外控倾向的形成可能有三种潜在来源：文化差异、社会经济水平的差异以及父母养育方式的差异。

有一项研究发现，在不同的文化间，个体的控制点倾向存在一些差异。在美国的一个相对封闭的社区中，主试比较了三个不同的群体，即印第安人后裔、墨西哥裔美国人和白种人。结果发现，一般而言，那些具有印第安血统的个体最具外控倾向，而白种人最具内控倾向，墨西哥裔美国人在I-E量表上的得分介于二者之间。以上结果与被试的社会经济水平无关，表明在控制点方面具有种族差异。

罗特还引用了一些早期实验的研究结果，这些结果表明，即便在某一特定的文化中，控制点的倾向也与个体的社会经济水平有关。社会经济水平较低的个体具有更大程度的外控倾向。

父母的养育方式被罗特视为是个体习得内、外控倾向的一个显而易见的原因。虽然他当时并没有拿出用以支持该论断的研究证据，但他认为父母在教养孩子时所采取的奖励和惩罚的方式若常与孩子的表现不一致，或是让人难以捉摸，那么这将会导致孩子形成外控的倾向（稍后我们将对此问题进行更详细的讨论）。

罗特总结了他的研究结果后指出，研究结果的一致性可以让我们得出如下结论：控制点是一种在各种情境中一贯发挥作用的可界定的个体特征。而且，个体的内外控维度对行为的影响是使不同的人在面对同一情境时采取不同的行动。此外，罗特声称控制点可以进行测量，而且I-E量表正是测量这一指标的有效工具。

罗特作出了如下假设，那些具有内控倾向的个体（如那些坚信自己能掌握自己命运的人）较那些具有外控倾向的个体更可能会：（a）为了改进自己将来在一些类似情境中的行为而尽力从该情境中获取更多信息；（b）更主动地改变和改善他们的生活状况；（c）更加看重成就和内在技能；（d）更能抵御他人的影响。

后续研究

自罗特编制了I-E量表以来，已有数百项研究检验了控制点与各种行为之间的关系。以下便是我们对这些涉及林林总总的人类行为的研究所作的简要介绍。

在1966年的论文中，罗特谈到了控制点与人体健康的关系。从那时起，许多其他的研究也做了相同的检验。在一篇对控制点研究的综述论文中，斯特里克兰（Strickland，1978）发现，内控型个体一般对自己的健康更加负责，且更可能采取健康的行为方式（如不吸烟以及养成更合理的营养搭配习惯等），而且在实际生活中更关心如何避免意外事故。顺便提及的是，研究已经发现内控型个体一般应激水平较低，而且很少患有与压力相关的疾病。

罗特关于教养方式与控制点存在相关的假设已经得到了部分验证。有研究表明，那些具有内控倾向的孩子的父母更有爱心，对孩子采取前后一致的、更加公正的奖惩约束，且更加注意教导孩子对自己的行为负责。具有外控倾向的孩子的父母，则更加专制和严厉，且不给孩子更多的自律空间（对这些发现的讨论参见Davis & Phares，1969）。

一项颇具吸引力的研究证明，控制点的概念有可能预示着社会性问题或灾难性问题。西姆斯和鲍曼（Sims & Baumann，1972）运用罗特的理论解释了为什么在亚拉巴马州死于龙卷风的人数会大大多于伊利诺伊州。这些研究者注意到：美国南部由龙卷风造成的死亡人数是中西部的5倍，进而他们着手追根溯源。他们一一排除了与物理因素有关的解

释，如龙卷风的强度及严重程度（事实上，发生在伊利诺伊州的龙卷风更为严重），龙卷风发生的时段（在这两个地区夜间发生的龙卷风次数相等）、商用和民用房屋的建筑类型（这两个地区都采用了相似的建筑技术）以及安全警报系统的质量（即使是在两个地区都没有安装警报系统之前，亚拉巴马州的死亡率也更高）。

排除了所有明显的环境因素之后，西姆斯和鲍曼认为差别可能源于心理变量，并提出把控制点概念列为一种可能的影响因素。他们采用I-E量表的修订版测试了伊利诺伊州和亚拉巴马州四个县的居民，这四个县经历过类似由龙卷风引起的死亡事件。他们发现，亚拉巴马州人的回答比伊利诺伊州人的回答具有更加明显的外控倾向。循着这一发现，并结合当地居民面临龙卷风时的行为及对他们对问卷中有关项目回答的分析，研究者得出了如下结论：内控倾向使个体在面临龙卷风时能存活下来（如内控者关注新闻广播或者会提醒其他人注意危险）。这是由于他们坚信自己的行为在改变事件的后果上能发挥作用。在该研究中，亚拉巴马人被视为“更缺乏自信，更缺乏采取有效行动的信心。……这些资料……为我们提供了颇具启发性的证据，那就是人格特征在决定人与自然相互作用的质量中发挥着多么积极的作用”（Sims & Baumann，1972年，P.1391）。

近期应用

自从该篇论文在1966年发表以来，如果说已有几百项研究借鉴了罗特的“控制点”理论，那么这一数字也可能大大低估了实际情况。事实上，可能有几千项这样的研究！罗特的理论受到了广泛的信赖，这就清楚地表明了人们对内、外控人格维度的影响及效度的广泛认可。下面是曾引用过罗特开创性工作的大量新近研究中的一部分代表。

通常，当我们讨论罗特对控制点的研究时，一个随之而来的话题便是被试的宗教信仰问题。很多有宗教信仰的人相信，有时将自己的命运交给上帝，不仅符合他们的心愿，而且是合适之举。然而，根据罗特的理论，此举将预示着一种具有消极意义的外控倾向。不久前，有人在《心理学与宗教》（Journal of Psychology and Religion）杂志上发表了一篇引人入胜的文章阐述这一论题（Welton，et al.，1996）。研究者使用多种控制点量表和分量表分别评估了被试的内控点水平、知觉到的受有影响力的其他人控制的水平、受运气和上帝控制的信念等。结果发现，在“受上帝控制”维度上得分较高的个体身上也发现了与内控型有关的一些优点。作者认为，如果一个人经罗特量表测量具有外控倾向，但这种外部力量被认为是一种对至高无上的神灵（如上帝）的强烈信仰，那么他也就不太可能陷入由外控倾向带来的一些典型问题（如无助感、抑郁、低水平的成就动机、缺少谋求改变现状的动机等）。

控制点的概念与个体对自己是否有选择权的感知有很大关联。事实

上，“有选择”会增加人们（尤其是内控型个体）的控制感。有一项研究综合了多种手段，包括动物研究、人类临床研究、神经成像（例如fMRI研究，见研究23）。该项研究的结果表明，人类对控制感的渴望不是后天习得的，而是一种进化得来的生存机制，是通过基因传递下来的（Leotti et al.，2010）。作者指出，如果没有“能做选择”（可产生最好结果的选择）的信念，我们将不会有任何动力去面对生活中的挑战，包括有助于使我们保持健康和远离危险的选择。这也就是为什么当你自由选择的权利（控制事情的能力）被剥夺时，常会产生一些病理性的结果，从明显的绝望到极端的生气和攻击。

另一方面，大量重要的跨文化研究在很大程度上基于罗特关于控制点这一人格维度的概念。例如，一项俄罗斯的研究考察了俄罗斯和美国大学生的控制点倾向以及他们对右翼独裁主义的态度（D’yakonova & Yurtaikin，2000）。结果显示，在美国学生中，较高水平的内控倾向与独裁主义有正相关，但在俄罗斯大学生中却并不存在这种关系。另一项借助罗特的I-E量表所进行的跨文化研究测查了在高度迷信的集体主义文化中癌症病人的心理调适状况（Sun & Stewar，2000）。有趣的是，该研究的结果表明，“即便是在一种具有超自然信仰的文化中，病人的内控倾向于与其进行的调适也有正相关，而外控倾向则与调适有负相关”（P.177）。引证罗特研究的领域除了上面已讨论过的这些之外，还包括创伤后应激障碍、控制与衰老问题、分娩方法、预期应激的应对方式、环境中噪音的影响、学术成就、白领犯罪、成年和儿童的酗酒问题、对儿童的性骚扰、自然灾害后的心理健康、避孕药的使用、人类免疫缺陷病毒（HIV）以及艾滋病的预防研究，等等。

结论

内、外控制点维度已作为人格的一个相对稳定的组成部分被人们广泛接受，它对预测众多情境中的行为起到了有价值的启示作用。之所以称之为“相对稳定”，是因为一个人的控制点在特定的环境中可以改变。那些具有外控倾向的人，由于在工作中被赋予了更大的权力和责任而常常向内控转变。而那些具有内控倾向的人在面对巨大压力和不确定性时，也可能会转向外控。而且，如果有机会，个体能够通过学习来提高内控水平。

隐含在罗特的控制点概念中的一个假设是，内控水平高的人们可以更好地自我调控并适应生活。尽管大多数研究都证明了这一假设，但罗特在其后期的著作中告诫我们说（参见罗特，1975），每个人特别是内控水平高的人，必须更加关注周围的环境。如果一个人打算改变一种不可能改变的情境的话，那么挫折、沮丧和抑郁很可能随之而来。当存在于个体以外的力量确实控制着行为的后果时，最现实也是最健康的行为方式是采取外控制点的态度。

Davis, W., & Phares, E.(1969).Parental antecedents of internal-external control of reinforcement.Psychological Reports, 24, 427-436.

D’yakonova, N., & Yurtaikin, V.(2000).An authoritarian personality in Russia and in the USA: Value orientation and locus of control.Voprosy Psikhologii, 4, 51-61.

Leotti, L., Iyengar, S., & Ochsner, K.(2010).Bron to choose: The origins and value of the need for control.Trends in Cognitive Science，14(10), 457-463

Rotter, J.(1975).Some problems and misconceptions related to the construct of internal versus external reinforcement.Journal of Consulting and Clinical Psychology, 43, 56-67.

Sims, J., & Baumann, D.(1972).The tornado threat: Coping styles in the North and South.Science，176，1386-1392.

Strickland, B.(1977).Internal-external control of reinforcement.In T.Blass (Ed.), Personality variables in social behavior.Hillsdale, NJ: Erlbaum.

Sun, L., & Stewart, S.(2000).Psychological adjustment to cancer in a collective culture.International Journal of Psychology, 35(5)，177-185.

Welton, G., Adkins, A., Ingle, S., & Dixon, W.(1996).God control—The 4th dimension.Journal of Psychology and Theology, 24(1)，13-25.

研究26 男性化、女性化……还是双性化

Bem, S.L.(1974).The measurement of psychological androgyny.Journal of Consulting and Clinical Psychology, 42，155-162.

你的性别是男是女？你是男人还是女人？你是男性化的还是女性化的？这是三个看似差不多的问题，答案却可能让你大吃一惊。对于第一个问题，答案一般是很明确的，因为这是基于染色体、荷尔蒙、生殖结构等做出的生物学回答。大多数人在回答第二个问题时也鲜有犹豫。实际上，你们从差不多四岁开始就都非常肯定自己的性别了。很可能你根本不用停下来花任何时间思考就知道自己是男人还是女人。

然而，第三个问题好像就没那么容易回答了。不同个体所拥有的“雄性”和“雌性”，或者说男性化和女性化特征的程度是不同的。想一下你认识的人，你可能会把他们中的某些人划到极度男性化一边（这些人应该都是男人），而某些人会属于极度女性化一边（这些人应该都是女人），但还有一些人的性别特征好像处于二者之间，他们既有男性化特征，又有女性化特征（这些人有男有女）。这些“分类”不带有任何的评判性，而仅仅是为了说明人类一个重要特征的差异性。这种男性化—女性化维度构成了心理学家通常提到的“性别”，而你对自己男性化程度和女性化程度的看法就是你的“性别认同”（gender identity）。性别认同是人格最重要的构成要素之一，它包括你自己以及他人关于“你是谁”的感知。

20世纪70年代以前，行为主义科学家（以及大多数非科学家）都持有一种两极化的性别观点：即人们的性别认同要么是彻底的男性化，要么是彻底的女性化，它们被看作是单一性别维度的两极。如果你基于这种观点进行性别认同测试的话，你的测验结果将处在单一维度上的某个位置，要么更接近男性化，要么更接近女性化。此外，当时的研究者和临床医师认为在一定程度上，心理调适与个人同其基于生物学性别的性别类型的契合度有关。换句话说，当时的观点是，为了达到心理健康的最佳状态，男人就应该尽可能男性化，女人就应该尽可能女性化。

然而，20世纪70年代早期，这种单维性别观点受到了挑战。安妮（Anne Constantinople，1973）在其发表的文章中指出：男性化和女性化并不是同一个维度的两极，相反地，它们应该成为衡量人类性别的两个独立的维度。也就是说，一个人拥有某种程度男性化特征的同时，也可能拥有某种程度的女性化特征。图7—1阐明了单维和二维性别观念的对比。

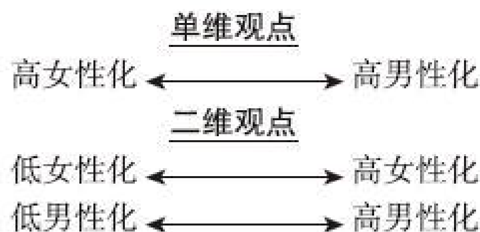


图7—1 传统单维性别模式与新近二维性别模式的对比

这一观点在你看来可能没什么值得大惊小怪，但在它最初提出的时候却可谓是一场革命。支持这种二维性别模式的代表人物是斯坦福大学的桑德拉·贝姆（Sandra Bem）。她向当时盛行的性别观念发起了挑战，当时人们认为健康的性别认同应该表现为一个人的行为符合其生物学性别的社会期望。而贝姆提出，一个更协调的人可以有效融合男性化和女性化两种行为，实际上比那些性别类型极度男性化或极度女性化的人更快乐、适应性更好。贝姆进一步发展了这项研究并开始研发一种基于二维性别模式的测量方法。在作为本文基础的这篇论文中，贝姆创造了“双性化”（androgynous）一词（“andro”意为男性，“gyn”意为女性），用来描述那些同时具有男性化和女性化特征的人，他们可以根据特定情境表现出最适宜的行为特征。此外，贝姆不仅认为某些人具有双性化的特征，而且当一个人从一种生活环境转移到另一种生活环境时，双性化的人格特征可以为更好的行为适应性提供有利条件。贝姆这样解释：

> 性别类型分明的人总想使自己的行为与已经内化的性别角色标准保持一致，并可能压抑任何与其性别不相符的行为来达成自己的目标。因此，一个自我概念中性别认同为男性化的人可能会抑制某些女性化行为，一个自我概念中性别认同为女性化的人也可能会抑制某些男性化的行为。而一个混合型的，或者说是双性化的人，就可以自由地做出男性化或女性化的行为了。（P.155）

举个例子，你可能认识一位优雅、敏感、轻声细语的女士（传统女性化特征），但她同时也是个野心勃勃、独立自主、身强体健的人（传统男性化特征）。另一种情况，你的一位男性朋友可能好胜、统治欲强、敢于冒险（男性化特征），但他有时也会表现出一些典型的女性化特征，如富有情感、有同情心、充满喜悦等。贝姆将这些人描述为“双性化”。本文解释了贝姆开发测量性别特征量表，即贝姆性别角色量表（BSRI）的理论基础和设计过程。

理论假设

每当科学家提出一种挑战传统观点的新理论时，他们也就必须承担起证明这种革命性观点正确性的责任。如果贝姆想要研究双性化的人与那些具有极端男性化或女性化特征的人之间的差异，她就需要找到一种

方法证明双性化特征的存在。换句话说，她必须测量出它。

贝姆认为要测量双性化特征，就需要有一种新的测量标准，它应该从根本上区别于先前所用的男性化—女性化量表。出于对这个目标的考虑，她所设计的量表有以下几方面的创新。

1.贝姆首先考虑的是所设计的性别量表应该抛开认为男性化和女性化是同一维度的两极的那种单维性别模式。她将两个独立的维度引入她的测验，分别用于测量男性化和女性化（见表7—2）。

2.她所设计量表的基础是男人和女人各自知觉到的社会期许的男性化和女性化特征。而先前那种测量方法是以在男人和女人身上可以普遍观察到的行为为基础，并非以美国社会期许作为判断标准。

如果一种特征对于男人来说更值得期许，就定义其为男性化特征；而一种性格对于女人来说更值得期许，那么就定义其为女性化特征。（P.155-156）

3.BSRI是为区分男性化、女性化和双性化而设计的。它主要是通过计算被试男性化和女性化得分之差来进行区分的。也就是用被试女性化得分减去其男性化得分，所得的差数就可以用来判断其男性化、女性化或双性化的程度了。

贝姆设计的测量标准是由一系列人格特征或特质组成的。为了得到一个性别分数，每种人格特征都在7点量表上进行评定，1到7点分别代表回答者所觉知到自身某种特质的程度。让我们看看这一量表是如何设计的。

方法

测试项目的选择

记住，贝姆要使用的是符合社会期许的性别特征。为了完成她的最终量表，她罗列出一长串她和她的心理学专业学生都认为有价值的特征。这些特征有代表男性化的，有代表女性化的，也有中性化的。三种特征列表均包括200个项目。她从斯坦福大学请了100名大学生（男、女生各一半），要求对于判断这些特征是否符合美国社会对于一个男人或者女人的期许，并在从1（完全不符合）至7（非常符合）的7点量表上进行评定。

按照学生们给出的分数，贝姆分别选取了两性评分最高的前20项性格特征作为男性化量表和女性化量表的测试项。同时，她也选择了20项可以忽略性别而每个人都期待拥有的性格（这些并不是双性化的测试项，仅仅用来反映中性性别特征），其中积极和消极的中性性格特征各

10项。在量表中加入这些测试项是为了保证不会由于测验题目全部是男性化和女性化的描述或者全部是社会期许的项目而过分影响被试的反应。最终的量表包含了这60个项目。表7—2呈现了最终选为BSRI项目的特质的范例。需要说明的是，在实际测量中，测试项不是按性别类型分开呈现的，而是按随机混合排列。

表7—2 修改后的性别角色量表

等级	女性化测试项	等级	男性化测试项	等级	中性测试项
___	有爱心	___	领导派头	___	适应性强
___	顺从	___	喜欢冒险	___	骄傲自满
___	充满喜悦	___	胸怀大志	___	捉摸不透
___	易被取悦	___	坚持己见	___	实事求是
___	有同情心	___	善于分析	___	效率低下
___	善解人意	___	个性鲜明	___	圆滑世故
___	优雅	___	果断	___	嫉妒心强
___	柔弱	___	自我满足	___	诚实可信
___	喜爱孩子	___	阳刚	___	喜怒无常
___	轻声细语	___	独立	___	值得信赖

（改编自表1，P.156）

按照如下标准为所有测试项评分：

1＝完全或几乎不符合

2＝大多数情况下不符合

3＝有时但很少符合

4＝偶尔符合

5＝常常符合

6＝大多数情况下符合

7＝完全或几乎符合

评分

女性化得分：女性化测试项分数总和÷10＝___

男性化得分：男性化测试项分数总和÷10＝___

双性化得分：女性化得分减去男性化得分＝___

说明：

女性化＝1.00或更高

接近女性化＝0.50至0.99

双性化＝-0.50至＋0.49

接近男性化＝-1.00至-0.49

男性化＝-1.00或更低

评分

正如前面所提到的，一个人进行BSRI测量，只需要用7点量表来衡量所有测试项与其自身的相符程度。相应的标准如下：1＝完全或几乎不符合；2＝大多数情况下不符合；3＝有时但很少符合；4＝偶尔符合；5＝常常符合；6＝大多数情况下符合；7＝完全或几乎符合。测试完成后，被试会得到三个分数：男性化分数、女性化分数和对本文而言最为重要的双性化分数。男性化分数是把所有男性化测试项得分相加再除以10得到的平均数，女性化分数也是按这种方法得到的，因此得分均在1.0至7.0之间。那么，你想到如何利用这两个分数算出双性化得分了吗？记住，这个量表可以分别测量男性化和女性化，但却不包含双性化的测试项。如果你想到双性化分数可以用女性化分数与男性化分数的差来表示，那么你就是正确的，因为贝姆正是这么做的。一个人的双性化分数就是用其女性化分数减去其男性化分数而得到的，其范围从-6至＋6。真的很简单。这里有三个典型的例子，他们分别代表了男性化、女性化和双性化性别类型的人。

詹妮弗的男性化分数是1.5，女性化分数是6.4。用6.4减去1.5得出她的双性化分数是4.9。理查德的男性化分数是5.8，女性化分数是2.1，因此他的双性化分数是-3.7。黛娜的男性化和女性化的分数分别为3.9和4.3，其双性化分数为0.4。

詹妮弗：女性化分数 = 6.4

—男性化分数 = — 1.5

双性化分数 = 4.90

理查德：女性化分数 = 2.1

—男性化分数 = — 5.8

双性化分数 = -3.70

黛　娜：女性化分数 = 4.3

—男性化分数 = — 3.9

双性化分数 = 0.40

看一下这些数据，我们所举的三个例子中，哪个人双性化程度最高呢？答案是黛娜。因为黛娜的男性化分数与女性化分数基本相同（双性化分数接近为0），并没有像詹妮弗和理查德那样显示出明显的性别取向。因此，黛娜的测量结果反映了其在自我觉知中缺乏明显的性别定型，更多的是一种男性化与女性化特征的平衡，这也恰恰是“双性化”一词的含义。

BSRI的测量结果可以这样解释：分数越接近于0（无论正负），表示双性化程度越高；在正方向上距0越远，表示女性化程度越高；在负方向上距0越远，表示男性化程度越高。

你也许想为自己测量一下。当然，就现在来说，你并不会得到一个理想的结果，因为你对这个量表的工作原理了解得太多了！你会把男性化、女性化和中性特质分开评分，而不是像在实际测量中那样混合在一起。不过，记住了这些告诫，你大可以试着测量一下。表7—2提供了简化的评分和解释说明。

结果

任何测量根据都必须具有信度和效度。“信度”是指测量的一致性，也就是各测试项对被试同一种特性反映的一致程度，以及经过重复测量得出一致结果的能力。“效度”是指实际测量的内容与想要测量内容的相

符程度——就BSRI来说，我们想要测量的内容是男性化特征和女性化特征。

BSRI的信度

对学生样本测量结果的统计表明，BSRI对男性化和女性化测量的一致性信度非常高。这就意味着男性化的20个测试项测量的都是同一种特质（假定为男性特征），女性化的20个测试项测量的也都是同一种特性（假定为女性特征）。为了确定测量的一致性信度不受时间影响，贝姆四周后又对最初测量过的60人实施了第二次测量。两次测量的结果相关程度很高，因此具有很高的“重测”信度。

BSRI的效度

为了证明BSRI的效度，必须首先对男性化与女性化测试项进行分析，以确保不是对同一特质的测量，这一点是非常重要的。因为贝姆所做研究的理论基础正是男性化和女性化是两个相对独立的维度，并且这两个独立的维度可以分开测量。贝姆证明这一点的方法是考查BSRI中男性化分量表和女性化分量表得分的相关系数，结果显示这两个分量表的得分无显著相关，功能各自独立。

接下来，贝姆需要证明这个量表测量的确实是男性化和女性化的性别特征。为了确认这一点，她分别分析了男人和女人在男性化和女性化分量表上的平均分。你可以预期，结果应该是男人在男性化项目上的得分更高，女人在女性化项目上的得分更高。这正是贝姆在大学生被试身上发现的，而且这一差异达到了很高的统计显著性。

最后，贝姆把被试按照前面讨论过的形式进行了性别归类，即男性化、女性化和双性化。她发现大多数人的女性化得分和男性化得分的差异都很小。换句话说，大多数被试是双性化的。表7—3呈现了在贝姆的研究中，男性化、女性化和双性化被试所占的百分比。

表7—3 被试中男性化、女性化和双性化的百分比

类型	男性	女性
女性化	7%	35%
接近女性化	6%	17%
双性化	35%	29%
接近男性化	19%	11%
男性化	33%	8%

注：被试数量＝917。

资料来源：改编自表7，P.161（合并了样本）。

讨论

贝姆文章中的讨论部分短小简洁，具有很强的说服力。转述它的最好方法恐怕就是直接完整地从原文中引用：

> 希望BSRI的出现和发展，能够鼓励性别差异和性别角色领域的研究者们去质疑那种只有性别角色典型化的人们才心理健康的传统观点，并转而关注这种更为灵活的性别角色观念对人类行为以及社会所产生的影响。当性别角色的硬性划分已经不再适用于这个社会时，也许双性化人这一概念会为心理健康建立一条更为人性化的标准。（P.162）

这段话说明了贝姆的研究是如何改变心理学的。在其文章发表几十年后，西方文化逐渐开始认同“有些人是双性化的”，而且同时拥有男性化特征和女性化特征不仅不被人排斥，反而被认为是一种优势。更多的男人和女人开始从事或参加那些在传统观念中被视为“仅限于”异性的职业、业余爱好、体育活动和家庭活动。从女CEO到“家庭主夫”，从女消防员、女军人到男护士、男教师，从女性掌管权力到男性展现其敏感的一面，你会看到，性别角色和性别期望的转变在社会中无处不在。

然而，这绝不是说文化已经对性别差异“视而不见”了。相反，对性别角色的种种期望仍会使我们对行为和态度的选择产生很大影响，而且性别歧视也依然是一个严重的社会问题。总的来说，男性还是被认为处事果断，而女性则常常被认为感情用事；绝大多数飞行员都是男性（96%），而几乎所有牙齿口腔保健医生都是女性（98%）。但是我们应该注意到，这种伴随性别分界的文化差异在不断减小，并将持续下去。

在贝姆这一新的性别观念的影响下，大量研究开展了起来。正如先前讨论过的，20世纪70年代以前，人们普遍认为，只有“社会性别与生物性别相匹配”才能最好地适应生活。也就是说，男孩和男人应该表现出男性化的态度和行为，女孩和女人应该表现出女性化的态度和行为。然而，双性化的发现改变了这一观点，研究者们也开始探索男性化、女性化和双性化个体之间的差异。

批评与后续研究

一项研究显示，双性化的儿童和成人趋于拥有更高水平的自尊，而且在多样化的情境下适应性更强（Taylor & Hall，1982）。另一项研究指出，双性化人群更易与异性建立亲密关系，可能是由于他们能更好地理解、接受彼此间的差异（Coleman & Ganong，1985）。更多近期研究揭示，具有积极的双性化特征的人在心理上更健康、更快乐（Woodhill & Samuels, 2003）。然而，随着时间的推移，贝姆等人发展的双性化理论也经历了许多改进和完善。

许多研究者指出：那些双性化程度高的人之所以体验到心理上有优势，可能更多的是因为他们表现出的是男性化特征，而非两种性别特征的平衡（Whitley，1983）。仔细想想，这是有道理的。很明显，许多传统的女性化特质，如依赖、自我批判、感情用事等都不符合社会期望。这就说明了为什么那些具有更多男性化特质的人会受到更好的对待，从而逐渐比别人的自信和自尊水平更高。但是，并不是所有男性化特质都是积极的，所有女性化特质都是消极的。每种性别所具有的特质都包括积极和消极两方面。

研究者们在此基础上进一步提出了更为精准的双性化概念，它包括四个维度：受赞许的女性化；不受赞许的女性化；受赞许的男性化；不受赞许的男性化（见Ricciardelli & Williams，1995）。像坚定、自信、坚强被看作是受赞许的男性化性格特征，而专横、吵闹、苛刻都是不受赞许的男性化性格特征。就女性化来讲，有耐心、敏感、责任感强都是受赞许的性格特征，反之像神经质、羞怯、软弱等都是不受赞许的性格特征。所以，通过一个人在上述各种特质的分布，就可以推断他所属的性别类型：积极男性化、消极男性化、积极女性化、消极女性化、积极双性化或消极双性化。

当性别特征根据积极和消极的性质进行更为细致的定义后，积极的双性化性格优势开始被推崇（例如Woodhill & Samuels，2003）。同时拥有最好的男性化和女性化品质的人，比那些仅拥有一种性别特征或结合了两种性别特征消极方面的人更全面、更快乐、更受欢迎、更招人喜爱、更灵活多变、也更爱自己。想象一下，把一个耐心、敏感、负责、坚定、自信、坚强（积极双性化性格）的人（男性或者女性）和一个神经质、羞怯、软弱、专横、吵闹、苛刻（消极双性化性格）的人进行对比，你就会更加清楚地理解这种对贝姆理论的修订。

贝姆在性别角色这一领域的研究一直处于领先地位。她应用其理论和研究对性别不平等问题进行辩论，并在其1994年出版的著作《性别透镜》（The Lense of Gender）中对此进行了详细的讨论。近期，她又开始关注婚姻、家庭以及儿童抚养的复杂问题，并出版了《一个不寻常的家庭》（An Unconventional Family，1998）一书。在书中，贝姆用自己和前夫康奈尔大学著名心理学家达里·贝姆（Daryl Bem）生活的经验，来探究一对夫妻如何避免形成传统的性别行为预期，如何真正平等地对待彼此，如何把孩子抚养成为“性别观念解放”且具有积极双性化性格的人。

近期应用

你在阅读本文时，可能会提出这样一个问题：用来测试男性化和女性化的那些测试项目是否一直有效呢？也就是它们是否可以一直用于区分男性化人群和女性化人群呢？事实上，你可能已经并不赞同其中的一

些测试项。毕竟这项研究是在几十年前进行的，而社会价值观对性别行为的期许也会随着时间而改变，对吗？这个问题的答案是“有可能！”20世纪90年代末开展的一项研究选取美国南部一所中等规模的大学的学生作为被试，对BSRI的所有项目进行重新评定。结果显示，除了其中两项，其他测试项仍能在很大程度上区分男性化和女性化（Holt & Ellis，1998）。那两项特例——“孩子气”和“忠诚”——在BSRI中都是对女性化的描述，但在1998年的研究中，它们并没有被评定为对于女性（而非男性）更受期许的特征。

然而另一项研究却得出了截然不同的结果。当选用美国东北部城市的一所大学的学生作为被试，对BSRI的测试项目进行验证时，结果出现了很大的差异（Konrad & Harris，2002）。研究发现：（1）女性被试认为，20个男性化测试项中仅一项（“阳刚”）对于男性来说更受赞许；（2）男性被试则认为，20个男性化项目中有13项受赞许的男性化性格特征；（3）女性被试只评出女性化项目中的2项（娇柔和轻声细语）对于女性来说更受赞许；（4）男性被试觉得有7项女性化项目对于女性来说更受赞许。

我们如何解释这两种研究结果之间的差异呢？一种可能是人们对性别的认同与地理区域有关。霍尔特和艾利斯的数据来自美国南部（还是一个相对较小的城镇），而康拉德和哈里斯的测试者则来自于美国东北部（而且是一个大型城市）。作者认为的另一种可能是：研究中的被试可能已经“猜到了”研究目的，从而使他们的答案带有了一定的倾向性：

> 明确地说，尽管要求被试对一种性别特征评分，但事实上仅仅指定性别作为测试目标这一点，就可能暗示被试该研究目的。既然有这种可能性，被试很可能会提供比真实情况更为平等的性别认同来表现积极的自我形象。（Konrad & Ellis，2002，P.270）

贝姆的工作对性征与性别的研究产生了巨大的影响，她的工作为与性别有关的研究奠定了基础。例如，有一项研究考察性别特征如何影响人们对处在领导地位的男女的看法（Ayman & Korabik，2010）。在诸多的影响因素中，研究者发现性别（部分地）决定了一个人是否能成为领导者。领导者大部分是男性，这是为什么呢？因为传统印象里，在成为领导的人身上，我们更多地看到男性化和社会主导性人格特征。另外，以任务为导向、更少表达情感的个体更有可能成为领导，而这些特质在男人身上更常见。然而，近期的研究表明，如果把群体的性别构成和特定的领导任务考虑在内，那么表现出双性化人格特征的个体更有可能成为领导者。这可能是由于双性化程度高的个体能更轻易地将任务导向和情感表达结合起来，而这正是使得许多领导工作更高效的关键。

结论

桑德拉·贝姆的这项研究之所以改变了心理学，是因为它改变了心理

学家、个体以及整个社会对人类一项最基本特征——性别认同的看法。贝姆的研究对开拓我们的视野具有重要的作用，使我们真正意识到什么是男人和女人，什么是男性化和女性化。也正是因为如此，我们每个人才有机会去拓展自己的行为模式、选择范围和生活目标。

Ayman, R.& Korabik, K.(2010).Leadership: why gender and culture matter.American Psychologist，157，157-170

Bem, S.L.(1993).The Lenses of Gender: Transforming the Debate on Sexual Inequality.New Haven, CT：Yale University Press.

Bem, S.L.(1998).An Unconventional Family.New Haven, CT：Yale University Press.

Constantinople, A.(1973).Masculinity-femininity: An exception to a famous dictum? Psychological Bulletin, 80, 389-407.

Holt, C., & Ellis, J.(1998).Assessing the current validity of the Bem Sex Role Inventory.Sex Role: A Journal of Research, 39, 929-941.

Konrad, A, & Harris, C.(2002).Desirability of the Bem Sex-Role Inventory for women and men: A comparison between African Americans and European Americans.Sex Role: A Journal of Research, 47, 259-272.

Ricciardelli, L., & Williams, R.(1995).Desirable and undesirable gender traits in three behavioral domains.Sex Roles, 33, 637-655.

Taylor, M., & Hall, J.(1982).Psychological androgyny: Theories, methods and conclusions.Psychological Bulletin, 92, 347-366.

Whitley, B.(1983).Sex role orientation and self esteem: A critical meta-analytic review.Journal of Personality and Social Psychology, 44, 773-786.

Woodhill, B., & Samuels, C.(2003).Positive and negative androgyny and their relationship with psychological health and well-being.Sex Roles, 48, 555-565.

研究27 和心脏赛跑

Friedman, M., & Rosenman, R.H.(1959).Association of specific overt behaviorpattern with blood and cardiovascular findings.Journal of the American Medical Association，169，1286-1296.

你是怎样的一个人？如果有人向你提出这一问题，你很可能会把自己较为明显的主导特征作为对该问题的回答。这些人格特征通常被称为“特质”，它们在使你成为一个独一无二的人这一方面起着重要的作用。人格特质具有跨情境和跨时间的一致性。支持人格特质理论的心理学家（并不是所有人都支持）认为，我们身上有很多种人格特质，例如双性化人格和控制点，这些特质在我们每个人身上的数量又有所不同。所有心理学家都感兴趣的是，人格特质使我们能够较稳定地预测出某人在某一特定情境中的所作所为。换句话说，人格特质理论家认为，对人格特质的剖面图进行分析，可以预测个体现在或将来的行为结果。因此，如果某种人格特质经研究能够预测你的健康状况，甚至能预测你死于某种心脏病的概率，那么人格心理学家对这种特质的兴趣急骤增加也就不足为奇了。

你也许知道有一种与健康有关的人格特征，它被人们称为A型人格。简单地说，A型人格是指一种特定的行为模式而不是个体的整个人格面貌。这种行为模式最早于20世纪50年代末期，由迈耶·弗里德曼（Meyer Friedman，1911~2001）和瑞·罗森曼（Ray Rosenman）两位心脏病学家首次提出。他们的理论和研究发现把心理学和健康联系在一起，并且对我们如何看待人格何以能导致或避免严重疾病产生了重大的影响。

理论假设

这两位医生研究构想的产生过程表明，对细微之处的仔细观察可能会引发科学上的重大突破。一次，弗里德曼医生请人为他候诊室里的家具重做皮套。做皮套的人说，沙发和椅子的皮套磨损方式很奇怪。具体说来，椅子的前沿比其余部分磨损得更快。这让弗里德曼医生想起似乎他的心脏病人常坐在椅子的边缘，这种现象促使他想知道他的病人（心脏病人）是否与其他科室的病人在某些重要人格特质上有所不同。

通过对公司高层和内科医生的调查，弗里德曼和罗森曼发现，人们普遍认为长时间处于疲劳驾车、面临最后期限、竞争性环境以及经济受挫等慢性应激状态下的人，更可能患上心脏病。于是，他们决定对这些

想法进行科学验证。

方法

利用他们的早期研究和临床观察，两位心脏病学家提出了一套反映一种特定的外显（可观察的）行为模式的模型，或者说特质组合。他们相信这一特定的行为模式与胆固醇水平的升高密切相关，并最终导致冠心病（CHD）。这种被称为A型行为的模式包括如下特征：（1）一种要达到个人目的的强烈而持久的内驱力；（2）在所有情况下都有一种热衷于竞争的倾向；（3）一种想得到别人的承认和不断超越自我的持久的欲望；（4）不断卷入有最后期限的事务之中；（5）习惯于风风火火地完成各种活动；（6）心理和生理上的过分敏感（P.1286）。

研究者紧接着提出了第二种被称为B型行为的外显行为模型，其特点与A型行为完全相反。这些行为的典型表现于不具备以下这些特征：内驱力、进取心、时间紧迫感、对竞争的渴望以及最后期限的压力。

弗里德曼和罗森曼接下来要做的就是寻找符合A型行为模式和B型行为模式的被试。为此，他们联系了各种各样的大公司大企业的经理和主管，并向他们解释这两种行为模式，请求经理和主管从自己的同事中挑选那些最符合该行为模式的人。最终选定的两组被试由各个级别的经理主管和普通职员组成，全部为男性。每组83人，其中A组的平均年龄为45岁，B组的平均年龄为43岁。所有被试都完成了与该研究的目的有关的一些测验。

首先，研究者用访谈的形式来了解被试父母的冠心病（CHD）的既往病史，被试自己心脏问题的既往病史，每周的工作、睡眠和锻炼的时间以及吸烟、喝酒和饮食习惯。在访谈过程中，研究者还要基于被试的肢体动作、谈话语气、紧咬牙关、体态、急躁神态以及被试关于内驱力、好胜心和时间紧迫感的自我报告等指标，确定被试是否具有其所在组（A型行为或B型行为）行为模式的全部特征。研究者最终确定A组的83名被试中有69人充分表现出A型行为模式，而B组的83名被试中有58人被判定为完全符合B型行为模式。

其次，要求所有的被试用日记方式记录自己在一周时间内的饮食情况。给每名被试分配一个编号，以便让他们在如实报告饮酒情况时不会感到为难。接下来，被试的饮食情况由医院中的一名营养学家进行归类和分析，该营养学家不了解被试的身份，也不知道被试属于哪个组。

再次，对每名被试进行抽血以测量他们的胆固醇水平和血凝时间。通过详细询问被试以往的冠状动脉健康状况以及标准的心电图数据来确定其冠心病的状况。这些心电图数据由罗森曼和另一名不参与该研究的心脏病学家分别进行独立解释和判断。除一例之外，他们的解释完全一

致。研究者还检查被试的眼睛确定患有角膜弓的被试数量。角膜弓又称老年环，指的是由血液里的脂肪堆积而导致的在眼角膜周围形成的一个不透明环状物。

现在，让我们概括一下弗里德曼和罗森曼的研究数据，看看他们究竟发现了什么。

结果

由访谈结果来看，两组被试的行为好像与研究人员所勾勒的两种行为模式的轮廓非常吻合。A组被试长期为承诺、抱负和内驱力所牵制。而且，他们显然渴望在所有活动中都力争上游，不管是专业活动还是娱乐活动。此外，他们也承认有强烈的获胜欲望。B组被试与A组被试相比有非常显著的不同，尤其表现在缺乏时间紧迫感上。B组被试似乎满足于他们的生活现状，而且不愿追求多个目标，并避免竞争性的情境。他们很少关心升迁的问题，而且往往将更多的时间用于与家人在一起以及参加非竞争性的娱乐活动。

表7—4概括了两组被试在测验和调查中最具可比性的特征。表7—5概括了有关血液水平和疾病情况的测量结果。正如表7—4所示，两组被试在几乎所有被测量的特征上都较为接近。尽管A组被试在多数测量指标上略高于B组，但在两者仅在每天吸烟的数量以及其父母患冠心病的概率方面存在显著性差异。

表7—4 A组和B组被试的个体特征的比较（平均值）

	体重	每周工作时数	每周锻炼时数	吸烟者人数	每日吸烟量	每日饮酒热量	饮酒总热量	脂肪热量	有子女的人数
A组	176	51	10	67	23	194	2 049	944	36
B组	172	45	7	56	15	149	2 134	978	27

（数据编自 pp.1289–1293）

表7—5A组和B组被试的血液和疾病的比较

	平均血凝时间（分钟）	平均血清胆固醇量	角膜弓（百分比）	冠心病（百分比）
A组	6.9	253	38	28
B组	7.0	215	11	4

（数据编自 P.1293）

然而，如果你浏览一下表7—5中两组被试的胆固醇情况和疾病情况，两组之间的确存在非常有说服力的差异。首先，尽管从表中的全部结果来看，两组被试的血凝时间似乎并不存在有意义的差别。血凝速度与心脏病及其他脉管疾病有关。血凝时间越慢，患病的危险性便越小。为了更严密地进行统计检验，弗里德曼和罗森曼比较了那些表现出典型

A型行为模式的被试的血凝时间（6.8分钟）和那些表现出典型B型行为模式的被试的血凝时间（7.2分钟）。结果发现，二者在血凝时间上的差异具有统计显著性。

表7—5中所示的其他发现是非常明确的。A组被试的胆固醇水平明显高于B组。如果将典型的A型被试和B型被试进行比较，这一差异则会更大。A型被试角膜弓的发生率是B型被试的三倍，而典型A型被试的角膜弓发生率则是典型B型被试的五倍。

使本研究得以载入史册的最重要的发现是两组被试在临床冠心病的发病率上存在极其显著的差异。A组中有23名被试（28%）出现了明显的冠心病发病迹象，而B组中只有3名被试（4%）有明显的迹象。当研究者对典型A型被试和典型B型被试进行进一步检验时，研究所得到的证据变得更为可靠。A组中有发病迹象的这23名被试全部具有典型的A型行为模式，而B组中的这3名有发病迹象的被试则是其中非典型B型行为模式的人。

对研究发现的讨论

作者隐含的结论是，A型行为模式是冠心病以及与此相关的血液异常的主要原因。不过，如果你仔细检查表中的数据，你会注意到对那些结果可以有其他的解释：一是A组的人报告其父母患有心脏病的人数更多。因此，或许是遗传而不是行为模式造成了两组之间存在的显著差异；另一引人注目的差异是A组被试每天的吸烟量更大。今天，一个众所周知的事实是吸烟可以引起冠心病。也许不是A型行为模式导致了这些结果，而是大量的吸烟引起了冠心病。

弗里德曼和罗森曼在他们对研究结果进行讨论时回答了上述两种潜在的质疑。首先，他们发现A组中吸烟少的被试（每天10支或更少）和吸烟多的被试（每天多于10支）患冠心病的人数一样多。其次，B组中吸烟多的被试多达46人，却只有两人有冠心病症状。作者根据这些发现认为，吸烟可能是A型行为模式的一个特征，但不是冠心病发病的一个直接原因。需要特别注意的是，该研究完成于40年前，当时人们对吸烟与冠心病之间的密切关系还没有达到像今天这样的认识。

关于父母的既往病史导致了这一差异的可能性，“研究数据显示，A组有心脏病家族史的30名被试中，只有8人（27%）患有心脏病，而其余没有心脏病家族史的53人中，却有15人（28%）患有心脏病；B组中有心脏病家族史的被试有23名，但其均未表现出临床的心脏病症状。”（P.1293）需注意的是，现在更多的对该因素进行严格控制的研究已经证明冠心病与家族因素有关。不过，我们还不清楚遗传的究竟是心脏病，还是某种行为模式（例如A型行为）。

研究发现的意义和后续发现

弗里德曼和罗森曼的研究对心理学研究的历史具有极其深远的影响。之所以这样说，主要是基于以下三个原因。

首先，该研究是最早系统探讨并提出个体的某些特殊行为模式能大大地增加某些严重生理疾病患病概率的研究之一。这一发现提醒医生，对成功的诊断、治疗、干预和预防而言，仅仅考虑到疾病的生理方面或许是远远不够的。

其次，该研究开创了探索行为与冠心病关系的一个新的研究方向，并且已产生了很多研究文章。A型人格及其与冠心病的关系已被提炼为如下要点：人们能够对心脏病高危人群的初次发病进行预防。

弗里德曼和罗森曼研究的第三个具有深远意义的影响是，它对于开创和发展健康心理学这一行为科学分支功不可没。健康心理学家从心理因素在健康的促进和维持、疾病的预防和治疗、致病原因及其在医疗保健系统中所处的地位来研究健康和医学的方方面面。

一项特别重要的后续研究值得在此作一介绍。1976年，罗森曼和弗里德曼发表了一项历时8年的研究结果。研究中的三千多名男性被试在该研究开始时被诊断为没有心脏病，且具有A型行为模式。这些人患心脏病的概率是B型行为模式被试的两倍，更易遭受心脏病的威胁，其报告的冠状动脉问题是B型被试的五倍。然而，更为重要的可能是，A型行为模式能够在独立于其他诸如年龄、胆固醇水平、血压或吸烟习惯等预测源的情况下，预测人们是否会患冠心病（Rosenman et al.，1976）。

现在你可能会问自己这样一个问题：“为什么A型行为模式会引起冠心病？”最为人们接受的一种理论回答是，A型行为者面对应激事件时，比非A型行为者的生理唤醒强度更高。这种极端唤醒状态致使身体产生过多的肾上腺素等激素，同时也提高了心率和血压。久而久之，这些对压力的过度反应会损害动脉，并进而导致心脏病（Matthews，1982）。

近期应用

弗里德曼和罗森曼或合作或独立地继续进行人格和行为变量在冠心病中所起作用的研究，并在该领域中扮演着带头人的角色。他们的研究连同其他研究已经发展出一种新的领域“心脏病心理学”（cardiopsychology），主要关注心理因素在冠心病的发病、病程、康复以及应对机制中的作用（Jordan，Barde，& Zeiher，2001）。本节讨论的是他们的第一篇论文，该文以及他们的其他许多的近期研究已被很多国家的众多研究所引证。A型行为的概念已经被完善、拓展和应

用到诸多研究领域。其中一些领域应用A型行为的概念是完全顺理成章的，而另一些领域或许会令你感到惊奇。

例如，有一项研究检验了A型行为与汽车驾驶之间的关系（Perry & Baldwin，2000）。得出的结论让人深信不疑，即“朋友们不会让具有A型行为的人驾车！”该研究发现，A型人格与驾驶事故有密切关联，产生更多交通事故、收到更多罚单、在路上更加不耐烦、表现出更多的公路暴怒，以及更多高风险的驾驶行为。看到这里，在你下次遇到交通事故之前，你或许想用本节结尾处的A型行为量表测试一下自己。

研究者们仍在继续探索人格因素与健康（尤其是冠心病）的关系，目前定义了一种新的典型行为模式：D型人格（Denollet et al.，2010）。字母D表示苦恼（distressed），D型人格的特点是大多数时间里都充满了负面情绪，对世界持有悲观的看法（“一切都糟透了”），害羞拘谨（与他人同处时感到不自在）。研究发现，这种行为模式与很多负面的健康事件有很高相关，包括动脉疾病、血管成形术、心脏搭桥术、心力衰竭、心脏移植、心脏病发作和因心脏问题造成的死亡等。研究者发现D型人格的个体处于社交环境中时会表现出明显的焦虑，同时在日常生活中会表现出较高水平的焦虑和抑郁。另外，他们还拒绝与他人谈论自己的心理不适感，因为他们非常在意周围人对自己的看法。有些人认为D型人格是A型人格的变体，有些人则认为它们是相互独立的。

有研究还发现A型人格（或许D型人格也是一样）会影响父母与青春期子女的关系（Forgays，1996）。在该研究中，研究者分析了900多名被试的A型人格特征和家庭教育环境。结果显示，如果父母具有A型行为，则其处于青春期的孩子也倾向于具有A型行为。尽管这并不令人吃惊，但却又一次引出了先天与后天的问题。孩子是继承了父辈A型行为的基因倾向，还是他们在A型父母对其养育的过程中习得了A型行为？福盖斯（Forgays）在他的研究中强调：“进一步的分析显示，能觉察到的家庭环境是青少年形成A型行为的一个独立的原因”（P.841）。不过，从最近的研究趋势来看，如果领养儿童研究、双生子研究和脑成像研究发现了A型/B型人格特征有基因或遗传的倾向，也并不令人感到吃惊（参见研究3中鲍查德针对基因影响人格所做的讨论和研究）。

结论

你有A型行为的人格特征吗？你是怎么知道的？正如本文开头部分所提及的你的内向或外向水平一样，A型/B型行为特征是人格的一部分。用以评估人们的A型或B型行为模式的测验已被编制出来。请参考下面的A型人格特征的条目，以大致了解你在多大程度上具有这些人格特征：

1.经常同时做几件事；

2.催促别人加快速度，让其快点把话说完；

3.遇到交通堵塞或排队等待时，显得很急躁；

4.讲话时有很多手势；

5.无事可做时很难坐得住；

6.讲话时很不耐烦，而且常常出言不逊；

7.比赛时总是想赢，甚至和儿童做游戏时也是如此；

8.在观看他人完成某一任务会变得不耐烦。

如果你怀疑自己是A型人格，你或许需要求助于一位有经验的医生或心理学家，以便做一个更仔细的评估。人们已经开发了一些打破A型行为与相关严重疾病关联的干预方案，很大一部分是在弗里德曼和罗森曼研究工作的基础上发展而来（例如George et al.，1998）。

Denollet, J., Schiffer, A., & Spek, V.(2010).General propensity to psychological distress affects cardiovascular outcomes.Evidence from research on the type D (distressed) personality profile.Circulation: Cardiovascular Quality and Outcomes, 3, 546-557

Forgays, D.(1996).The relationship between Type-A parenting and adolescent perceptions of family environment.Adolescence, 34(124), 841-862.

Gearge, I., Prasadaro, P., Kumaraiah, V., & Yavagal, S.(1998).Modification of Type A behavior pattern in coronary heart disease: A cognitive-behavioral intervention program.NIMHANS Journal，16(1), 29-35.

Jordan, J., Barde, B., & Zeiher, A.(2001).Cardiopsychology today.Herz, 26, 335-344.

Matthews, K.A.(1982).Psychological perspectives on the Type A behavior pattern.Psychological Bulletin, 91, 293-323.

Perry, A., & Baldwin, D.(2000).Further evidence of associations of Type A Personality scores and driving-related attitudes and behaviors.Perceptual and Motor Skills, 91(1)，147-154.

Rosenman, R.H., Brond, R., Sholtz, R., & Friedman, M.(1976).Multivariate prediction of CHD during 8.5-year follow-up in the Western Collaborative Group Study.American Journal of Cardiology, 37, 903-910.

研究28 个人与集体

Triandis, H., Bontempo, R., Villareal, M., Asai, M., & Lucca, N.(1988).Individualism and collectivism: Cross-cultural perspective on self-ingroup relationships.Journal of Personality and Social Psychology, 54, 323-338.

如果说，有哪一种人类的特征能几乎被所有的心理学家所认同的话，那就是人类的行为从来不会在真空中发生。即使是那些最强调内部动机、个性需求和遗传驱动力的心理学家们也会考虑把各种各样的外部环境力量引入最终导致个体行为和特质形成的方程式中。在过去的三四十年间，心理学领域已经开始渐渐认识到文化是一种对人类具有重大影响的环境因素。事实上，研究人员发现，很少有在所有或者在大多数文化中具有一致性和稳定性的人类行为模式（跨文化一致性的扩展分析参见研究22中关于埃克曼的面部表情研究的讨论）。这在涉及人类的相互作用和人际关系的行为领域中显得尤为突出。人际吸引、性、感动、个人空间、友谊、家庭动力、父母养育方式、儿童行为期望、求偶仪式、结婚、离婚、合作与竞争、犯罪、爱和恨等都深受文化因素的影响。可以肯定地说，如果不考虑一个人的文化背景，那就根本无法全面或准确地了解这个人。

从概念上讲，文化是很容易理解的，但具体到实践领域，文化却成了一块难啃的骨头。试想，你如何能够理清那些交织在一起的、共同塑造了现在的你的所有文化因素呢？大多数文化是如此复杂，以至于无法得出有据可循的结论。例如，结肠癌在日本的发病率比在美国低得多。日本和美国是两种不同的文化，那么到底哪些文化因素可以解释发病率的差异呢？是对鱼的需求量的不同？是消费大米数量的不同？是饮酒量的不同？是生活节奏和应激水平上的差异？或许是两国宗教习惯的差异影响了健康？来自亲戚朋友的支持的差异是否会影响健康状况？抑或更可能是以上两个、三个或者所有这些因素连同其他因素共同导致了这一结果？问题的关键是，如果你打算把文化列为全面了解人类本性的因素之一，那就需要用可靠和有效的方式来定义文化差异。正是在这一点上，哈里·川迪斯在现代心理学史上崭露头角。

从20世纪60年代起，川迪斯在芝加哥大学俄巴娜—香槟分校心理学系致力于揭示和提炼文化及其背景下的人们的基本特性，以便以有意义的方式对这些文化加以区分和研究。这项工作贯穿于川迪斯的整个事业生涯。发表于1988年的这篇论文解释并论证了他对跨文化心理学所作出的巨大贡献，它阐述了个人主义文化和集体主义文化。今天，这一基本的文化差异维度构成了每年的心理学、社会学以及很多其他领域数百项研究的基础。在这篇论文中，川迪斯认为，在某种程度上被界定为个人主义或集体主义的特定文化，正以一种复杂和广泛的方式决定着其成员

的行为和人格。

用通俗的话讲，集体主义文化是这样的一种文化，即生活在该文化中的个体的需求、欲望、成就都必须从属于他所属的内群体（ingroup）的需求、欲望以及目标。内群体可以是一个家庭、一个部落、一个村庄、一个专业组织或者甚至整个国家，具体视情况而定。在这些文化中，个体的大部分行为的产生取决于该行为是否有利于个体所属的更大群体的整体利益，而不是该行为是否能为个体提供最大限度的个人成就。人们所归属的内群体往往会保持长期稳定，而且个体对其所属群体的承诺通常都很高，甚至当一个人在群体中的地位对其而言已变得困难或者不愉快时也是如此。个体依靠他们所属的群体来帮助他们满足自己的情绪、心理和实际的需要。

相反地，个人主义文化更看重个体的幸福和成就，而不是其所属群体的需要和目标。在这些文化中，群体对其所有成员的影响可能是微乎其微的。个体对其所属群体没有太多情感上的依恋，并且如果它对个体的要求过于苛刻，个体则宁愿脱离该群体而加入或者重新组建一个新的群体。在个体主义文化中，个体对群体的承诺很少，所以对个人而言，同时是多个群体的成员是很常见的事情。然而，没有哪个群体能单独对其行为施加太多的影响。在该篇论文中，川迪斯及其来自不同文化背景的合作者，描述了大量典型集体主义文化和个人主义文化特征。所有这些都被总结在表7—6中。当然，这些特点是高度概括化的，因此不管是在集体主义文化还是在个人主义文化之中，总有一些例外的情况。

按照川迪斯的观点，个人主义文化集中在北欧、西欧和那些历史上受北欧人影响的国家。此外，典型的个人主义文化有一些共同的特征：具有边界，大量的外来移民，社会和地理上的迅速变迁。“所有这些特点倾向于使群体的控制力变得不太确定。在美国、澳大利亚和加拿大，高度的个人主义是与上述特点相一致的。”（P.324）他认为，世界上其他大多数地区属于集体主义文化。

理论假设

川迪斯在其文章的开头部分指出：

> 文化是一种模糊的结构体。如果我们要了解文化和社会心理现象的联系，我们必须通过确定文化差异的维度来分析它。最有前景的维度之一就是“个人主义—集体主义”。（P.323）

在川迪斯的很多研究和著作中所隐含的研究假设是，当我们按照“个人主义—集体主义”模型来界定和诠释文化时，这一模型能够解释我们所见到的表现在人类行为、社会交往和人格中的大部分差异。在该文中，川迪斯试图概括出其理论的广泛的潜在用途（参见表7—6），并报告了他为了验证和阐明其个人主义—集体主义理论所做的三项科学研究。

方法

如前所述，该篇论文报告了三项独立的研究：第一项研究的被试全是美国人，以便更加清晰地定义适用于美国的个人主义的概念；第二项研究的目的是比较个人主义文化（美国）和集体主义文化（如日本和波多黎各），其核心在于比较在两种文化类型中个体与他们所属群体之间的关系；第三项研究用以检验如下假设，即集体主义文化中的成员觉察到自己获得更好的社会支持，且一贯享有令人满意的人际关系，而个人主义文化中的成员则报告他们常常感到孤独。所有这些数据都是通过问卷调查的形式来收集的。下面我们将简要介绍每项研究及其发现。

表7—6 集体主义文化和个人主义文化的差异

集体主义文化	个人主义文化
· 牺牲：把群体目标置于个人目标之上	· 享乐主义：关注令个人满意的目标胜过群体目标
· 把自我看成是所属群体的延伸部分	· 认为自我与所属群体相独立
· 关心群体是最重要的	· 自力更生是最重要的
· 因对群体所作的贡献而得到回报	· 因个人的成就而得到回报
· 个人和文化方面的单一性	· 个人和文化方面的丰富化
· 高度服从明确的群体规范	· 不太服从群体规范
· 更崇尚爱、地位和贡献	· 更崇尚金钱和财产
· 多与群体内的成员合作，少与群体外的成员合作	· 经常与群体内的成员以及群体外的成员合作
· 遵从“纵向的人际关系”（孩子与父母、雇主与雇员）	· 遵从“横向的人际关系”（朋友与朋友、丈夫与妻子）

续前表

集体主义文化	个人主义文化
· 通过频繁地询问及干预孩子的个人生活来养育孩子	· 通过与孩子保持距离，使孩子独立及保护孩子的隐私来养育孩子
· 在达到目标时更多以人为定向	· 在达到的目标时更多以任务为定向
· 宁愿隐瞒人际冲突	· 宁愿面对人际冲突
· 个人对群体所承担的责任很多，但群体会以高水平的社会支持、资源和安全感作为回报	· 对群体的责任较少，个人权利较多，但从群体中得到的支持、资源和安全感较少
· 朋友不多，但朋友间的感情较深厚，关系稳固持久并承载着很多责任	· 交朋友容易，但其中多是不那么亲密的熟人关系
· 内群体较少，而且把所有的其他人视为一个大的外群体的成员	· 有很多内群体，但较少把所有其他人看成是外群体的成员
· 和所属群体内的成员关系非常融洽，但可能与群体外的成员有较大的冲突	· 内群体规模较大，且很可能发生群体内部的人际冲突
· 惩罚多以羞耻感（外部的）的形式出现	· 惩罚多以罪恶感（内部的）的形式出现
· 经济发展速度较慢，工业化水平低	· 经济发展速度较快，工业化水平高
· 社会弊端较少（犯罪、自杀、虐待儿童、家庭暴力、心理疾病）	· 各种社会弊端较多
· 疾病发生率低	· 疾病发生率高
· 婚姻更美满，离婚率低	· 婚姻不太幸福，离婚率较高
· 竞争较少	· 竞争较多
· 关注家庭等小群体的利害得失，而不是更大的公共利益	· 更大程度地关注公共利益

（引自 Triandis，1988，pp.323-335）

研究1

第一项研究的被试是川迪斯所在的芝加哥大学心理系的300名在校大学生。研究者要求每位学生完成一份包括158个项目的问卷，用以测量他们有关“个人主义—集体主义”的行为和信念倾向。如果被试赞同“只有那些依靠自己的人才能在生活中获得成功”，那么他便持有个人主义的立场。而被试如果同意“当我的同事把他们的隐私告诉我时，我们的关系更加亲密了”，则说明该被试持有集体主义的观点。此外，问卷中包括5个生活情景故事（Scenarios），这些情景故事把被试置于一种假设的社会情境中，并要求他们预测自己的行为。该论文中所提供的一个例子是，让被试想象自己想要做一次群体内其他成员都反对的长途旅行。问被试在决定是否去旅行时，会在多大程度上考虑父母、配偶、亲戚、密友、熟人、邻居以及同事的意见。

对回收的问卷进行分析后发现，在被试的回答中几乎有50%的变异能够被三个因素所解释，这三个因素是：自力更生、竞争、与所在群体的亲疏关系。只有14%的变异能被“对群体的关注”这一因素所解释。川迪斯对研究1的结果做了如下概括：

这些数据表明，美国式的个人主义是一个包括多方面含义的概念。具体包括：关

心自己的目标胜过群体的目标，不太关注群体的观点；自立与竞争相结合，与群体有些疏远；做决定时根据的是自己的看法而不是其他人的观点，以及对群体不太关心（P.331）。

他还指出，该问卷所含的项目和生活情景故事能有效地测量美国文化的个人主义程度，但川迪斯对这些项目在其他的文化背景中是否同样有效尚存疑问。

研究2

在该研究中，研究所提出的问题是："集体主义文化中的人们是否更乐意使自己的个人需要服从于集体的需要？"被试是91名芝加哥大学的学生、97名波多黎各学生、150名日本大学生和106名年长的日本人。测量集体主义特征的问卷由144个项目组成。问卷被译成西班牙语和日语后交由所有被试进行回答。经之前的研究验证，该问卷的题目分别测查关于集体主义的三个倾向：关心群体、自我与群体的亲密关系和使自己的目标服从于群体的目标。

该项研究的结果互不一致且很有趣：一些结果支持"个人主义—集体主义"的理论，而另一些则似乎又对它予以反驳。例如，日本学生比伊利诺伊的学生更在意同学和朋友的观点，但波多黎各学生并没有表现出这一特点。此外，尽管日本被试在群体获得荣耀表达了自豪感，但他们在生活中只关注他们所在的某些群体的观点，并牺牲自己的个人目标以服从这些特定群体目标，而不是对所有群体都如此。尽管服从是集体主义文化的一个共同特征，但事实上日本被试比美国被试更少地表现出对群体的服从。有结果表明，随着集体主义文化变得更加丰富化和西方化，它们可能正在经历一个向个人主义转变的过程。作为这一观点的又一证据，研究发现年长的日本被试比日本大学生倾向于认同其所属的群体。

在这一点上，你或许会问，第二项研究的结果如何纳入川迪斯的理论呢？川迪斯把它们解释成一种警示，即有关集体主义文化和个人主义文化的结论不应该过分概括化，而必须仔细分析并有选择地应用于特定的行为、情境和文化之中。他对这一观点做了如下阐述：

> 该项研究的数据告诉我们要限制和细化我们对集体主义的定义，……我们必须考虑到每个独立的社会行为领域。被界定为服从所在群体的规范、需求、观点，而且情绪上对群体有依恋感的集体主义对于特定的群体和社会行为领域而言是一个非常具体的概念。……集体主义在每种特定的文化中分别呈现出不同的形式。（P.334）

研究3

第三项研究正是试图要完成川迪斯在上述引语中所提出的研究任务，对研究的核心加以限制和细化。该研究拓展了先前的研究结果，即集体主义社会为它们的成员提供了高水平的社会支持，而个人主义文化中的个体则在更大程度上体验到孤独。该研究所采用的集体主义—个体

主义问卷由72个项目组成，总共施测了来自芝加哥大学和波多黎各大学的100名被试，男女各半。被试还填写了另一份问卷以测量他们所知觉到的社会支持水平和孤独程度。

该研究的结果清楚地表明，集体主义与社会支持存在正相关，也就是说，随着集体主义程度的提高，相应的社会支持水平也会提高。而且集体主义与孤独程度呈负相关，即随着集体主义作用的提高，被试所意识到的孤独程度降低。作为川迪斯理论模型的进一步的证据，对该研究中的美国学生来说，最重要的因素（能解释最大的变异）是“在竞争中自力更生”，而对波多黎各学生而言，最重要的因素是“合群”（与他人的互动）。这些结果正是人们可以由“个人主义—集体主义”理论预期得到的。

讨论

川迪斯解释说，总而言之，本文中所述及的这些研究既支持又改进了他对集体主义和个人主义的定义。回过头来看表7—6中所列出的两种文化类型的特征，其所显示的状况似乎是相互对立、非此即彼的，即个人主义文化和集体主义文化看来几乎是完全对立的。不过，本文证明了这两种文化类型似乎处于一个连续体的两端，而某一特定社会则处于两端之间的某个点上，这个点通常更接近于连续体的一端而远离另一端。此外，具体到某一文化，其中特定的个体、群体、亚文化以及情境可能会偏离该文化在连续体上的大体位置而向某一端移动。图7—2呈现了这一解释的图示。川迪斯指出：“简言之，这些实证研究表明，我们必须把个人主义和集体主义看作是一种多维的结构，……每一维度的性质和特点主要取决于其所属群体的现状、情境线索以及所研究的具体行为。”（P.336）

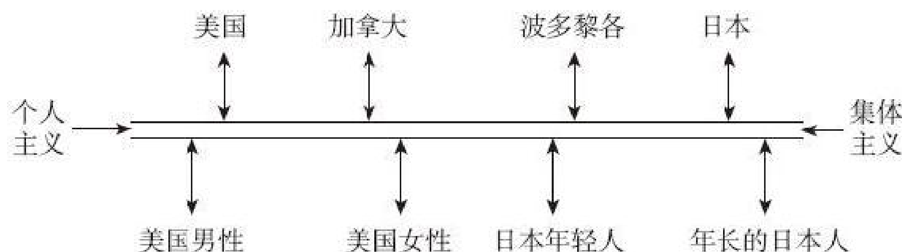


图7—2 集体主义—个人主义的文化连续体（文化和亚文化的近似位置）

研究发现的意义和相关研究

尽管该研究的历史很短，川迪斯的研究结果已经成为心理学家看待

人类行为的基础之一。在你打开大多数心理学分支——基础心理学、社会心理学、发展心理学、人格心理学、人类性行为、变态心理学和认知心理学等的新版教科书时，你将很难找到一本不涉及川迪斯的这项研究，或是他做的很多其他有关“个人主义—集体主义”的研究。可以说，在有关文化对于决定人格和社会行为所起作用的当代研究中，由川迪斯所阐述、澄清和提炼的个人主义—集体主义文化维度是最可靠、最有效和最具影响力的因素。此外，该维度能被应用于非常广泛的研究领域。以下正是这方面的两个应用实例。

在川迪斯的这篇论文中，他所提供的证据表明，集体主义和个人主义这两个心理社会概念可能会对一种特定文化的成员的身体健康产生重大影响。这项与冠心病有关的研究就是一个很好的例子。一般而言，在集体主义社会中，心脏病的发病率比在个人主义社会中要低。川迪斯认为，与心脏病密切相关的令人不快和紧张的生活状态在个人主义文化中更为常见。为了竞争和实现个人目标，孤军奋战的个体承受着巨大的精神压力。除了这些消极的生活事件之外，个人主义的社会结构本身就缺乏对其成员提供的社会凝聚力和支持力量，而这种力量却恰恰能够减少压力给人体健康所带来的影响。当然，正如本文开头部分所讨论的那样，有很多因素可以解释在心脏病以及其他疾病的发病率方面出现的文化差异。但很多研究已经表明，集体主义文化中的成员迁居到个人主义文化的国家后，也越来越易于患上包括心脏病在内的各种疾病。

或许，更有说服力的是在同一种文化中的两个不同的亚群体之间所存在的差异。正像川迪斯所指出的（P.327），研究人员做了一项以3000名日裔美国被试的研究，主要是比较那些在生活方式和态度上已完全适应了美国风格的日裔美国人和那些仍然保持传统日本生活方式的日裔美国人。结果发现，在排除了胆固醇水平、锻炼、吸烟、体重等因素的情况下，前者心脏病发病率仍然是后者的5倍。

当然，你也许已料到“集体主义—个人主义”维度会影响到特定的文化群体之间在养育子女的方式上的不同，事实上的确如此。集体主义社会中的父母更加注重培养儿童的“集体自我”（collective self），这种集体自我的特征是对群体规范的服从，对群体中权威的服从，行为具有跨时间和跨情境的可靠性和一致性。家长常常以公开或微妙的方式对做出符合群体目标的行为模式的儿童给予奖赏（Triandis，1989）。在这种情形下，只是因为自己不喜欢而拒绝从事群体期望自己做的事情是不能被接受的，这种情况也极少发生。然而在像美国这样的高度个人主义的文化中，这种拒绝是一种正常的反应，而且常常受到人们的重视和尊敬！这是因为在个人主义文化中，父母重视培养儿童的“个人自我”（private self）。这就意味着儿童会因具有自力更生、独立、自知之明以及最大限度地发挥自己潜能的行为和态度而得到奖赏。另一个看待这种差异的角度是，在个人主义文化中，反抗（在社会可接受的范围内）和独立倾向被看成是一种人格资产（personality assets），而在集体主义社会里，它们被看成是负债（liabilities）。文化经由父母向儿童传达的关于资产或负债的信息是强大而明确的，它对儿童的成长有着强

有力的影响。

近期应用

川迪斯的工作影响着广泛的研究领域。其中一篇文章将川迪斯的理论应用到研究两种文化条件下大学生球迷的态度问题（Snibbe et al.，2003）。研究者对观看重要足球比赛的美国（玫瑰杯美式足球赛）和日本（闪光杯足球赛）的大学生进行问卷调查，要求这些大学生在足球比赛前和比赛后，分别对自己球队与对方球队的学校和学生进行评定。在这两场比赛中，都是具有优良学术氛围的大学输掉比赛。然而，在两种不同的文化背景下，学生的反应差别很大："美国大学生无论是在赛前还是赛后，均表现出对自己球队更加积极的态度。相比之下，日本大学生的评分没有表现出对本队的偏爱。而且，日本大学生对学校的评分反映了每一所大学在社会上的地位，对学生的评价则是根据目前的状态。"（P.581）

另一项研究采用川迪斯的理论模型对孤独感进行了跨文化研究（Rokach et al.，2002）。该研究选取了来自北美和西班牙的1000多名被试，要求他们完成关于孤独感成因的调查问卷，其中包括个人体缺陷、发展性障碍、缺乏亲密感、迁徙和分离，以及社会排斥感。"结果表明，孤独感的成因的确受到文化背景的影响。而且，北美被试在这五个维度上的得分较高。"（P.70，附加强调）

最后，有一项研究突出了川迪斯研究工作的一个很重要的方面。当我们在对集体主义和个人主义加以研究和比较时，这种比较绝不只限于在国家之间进行。很多国家本身就包括了多种水平的集体主义和个人主义的亚类型。在世界上没有哪个国家比美国拥有更多的不同形式的集体主义和个人主义。一项根据川迪斯的理论模型完成的一项有趣的研究（Vandello & Cohen，1999）以图表的形式描绘了美国集体主义和个人主义的分布情况。在继续往下阅读之前，请想一想你认为最典型的个人主义和集体主义会出现在什么地方。研究发现，美国的南方腹地具有突出的集体主义倾向，而位于大平原和落基山脉的地区则呈现更明显的个人主义倾向。不过，即便是在这些不同的地区内，也可以发现不同层次的个人主义和集体主义的亚文化群体。

结论

川迪斯给所有的社会科学提供了一种看待基本的文化差异的新视角。随着世界变得越来越小、多种社会文化日益交织，我们所直接经历的文化多样性常常会导致一些潜在的误解、沟通障碍、冲突和挫折。对于集体主义和个体主义文化差异的觉察和尊重，有可能让我们朝着缓和

文化分歧与促进世界和谐的积极目标迈出一小步，但却是富有意义的一步。

Rokach, A., Orzeck, T., Moya, M., & Exposido, F.(2002).Causes of loneliness in North American and Spain.European Psychologist, 7, 70-79.harmony

Snibbe, A., Kitayama, S., Markus, H.& Suzuki, T.(2003).They saw a game: A Japanese and American (football)field study.Journal of Cross-Cultural Psychology, 34, 581-595.

Triandis, H.(1989).The self and social behavior in differing cultural contexts.Psychological Review, 96(3), 506-520.

Vandello, J., & Cohen, D.(1999).Patterns of individualism and collectivism across the United States.Journal of Personality and Social Psychology, 77(2), 279-292.

第八章 心理病理学

大多数没有学过心理学的人会有这样的印象：这一学科主要涉及心理疾病的分析与治疗（心理学中研究精神病的一个分支被称为变态心理学）。然而，正如你所看到的，在这本书中，几乎所有的研究都把焦点集中在正常行为上。几乎所有的心理学家对正常行为的兴趣远远超过了对异常行为的兴趣，因为大量的人类行为并非是病态的，而是正常的。如果只研究占人类行为中很少比例的异常行为，我们就不能更多地了解人类的本性。尽管如此，对许多人来说，精神病也仍然是心理学中众多使人感兴趣的研究领域之一。在这里面包括了心理学史上一系列必不可少的研究。

第一项研究在精神卫生领域已经谈论了30多年。在这个研究中，有几个正常人假扮成精神病人进入精神病院，来了解医生和工作人员是否能把他们与真正的精神病人区别开。在所有关于心理学历史的书中都必须提及西格蒙德·弗洛伊德。因此，第二个研究是通过他的女儿安娜·弗洛伊德的著作来讨论弗洛伊德提出的“自我防御机制”这一经久不衰的概念。第三个研究是一个以狗为被试的实验，该研究阐释了一个被称为“习得性无助”的现象。这个研究与心理病理学的关联在于它引发了大量解释人类临床抑郁症的理论。第四个研究是一个极有趣的、非常有名的实验，它是关于过度拥挤的老鼠及由拥挤而导致的行为异常的研究，该实验可能对人类具有非常重要的启示作用。

研究29 究竟谁是疯子

Rosenhan, D.L (1973).Onbeingsaneininsane places.Science，179, 250-258

辨别“正常”行为和“异常”行为是心理学的基本任务。对异常行为的界定在决定一个人是否被诊断为精神病中起关键作用，并且诊断又在很大程度上决定病人接受的治疗。然而，区分正常与异常的界限并不像你所认为的那么清晰。所有的行为可以被看做是位于一个连续的轴上，正常或称为“有效的心理功能”位于一端，异常或“心理疾病”位于另一端。

在这个连续的轴上，决定人的一种特定行为在这个轴上所处的位置通常是精神卫生专业人员的责任。为了做这个决定，临床心理学家、精神病学家和其他行为科学家及临床医生可以应用下面的一条或多条标准。

·行为场合。这是一种主观判断，但你知道，某些行为在某个特定的情景中显然是古怪的，而在另一个特定情景中就没那么显眼。例如，你在院子里给草坪浇水没有什么古怪，而你在暴风雨中给草坪浇水就是行为古怪！判断行为古怪必须认真考虑行为发生的场合。

·行为持续存在。我们都曾有过疯狂的时刻。一个人有可能偶尔表现不正常行为，不一定证明他患了精神疾病。例如，当你正沿着一条繁华的商业街的人行道散步，你忽然收到了一些振奋人心的好消息，然后你忍不住跳起舞来，就这样经过了大约半个街区。尽管这种行为有些不正常，但并不表示你精神异常，除非你一直跳下去，每周或每天的大部分时间都是如此。因此，对精神疾病的这一诊断标准，要求一个古怪的、反社会的或失控的行为模式持续存在。

·社会偏差。当一个人的行为严重违反了社会期望和规范，这种行为就可能被认为是社会偏差。当偏差行为非常极端和持久，如幻听和幻视，就是精神病的证据。

·主观痛苦。我们通常知道自己的心理问题及其给我们带来的苦楚。当一个人是如此害怕封闭的环境，以至于他或她不能乘电梯，或者某些人发现自己不能与别人建立富有意义的人际关系，不需要专业人员的提醒，他们就清楚自己正处于心理痛苦之中。精神卫生专业人员做出心理诊断时，把主观痛苦看作是一种重要的指标。

·心理障碍。当一个人发现，由于心理问题他不能满意地生活

了，这就可以被认为患了心理障碍。例如，一个人害怕成功，因此，生活中的每一次新的努力都会在某种程度上遭到他所患心理障碍的破坏。

·对功能的影响。一个人的问题行为对其过自己想要的或社会可接受的生活能力的干扰程度，可能是心理诊断中最为重要的因素。某种行为可以是古怪的、持续的，但如果它没有伤害你的生活功能，就不代表真正的病理特征。例如，假设你有一个不能控制的需要，即每天晚上睡觉之前，站在你的床上唱国歌。这种行为当然是古怪的和持续的，但除非你把邻居吵醒或干扰了家庭其他成员，否则你的行为只对你个人的一般生活功能有很小的影响。因此，它不能被归类为临床问题。

精神病的这些症状和特点，都依赖于心理学家、精神病学家和其他心理卫生专业人员的判断。因此，前面的指导标准存在两个问题：心理卫生专业人员真能区别精神疾病和精神健康吗？区别错误的后果是什么？大卫·罗森汉（David Rosenhan）在其著名的精神病院研究中对此进行了探讨。

理论假设

罗森汉的问题是：决定观察者（即做诊断的那些人）做出心理诊断的，究竟是由于病人本身具有的特征，还是病人所处的环境或场合。他推想，如果精神疾病诊断标准的制定和精神卫生专业人员所接受的诊断训练是恰当的，那么，专业人员应该能鉴别心智健全和心智不全（专业名词“心智健全”和“心智不全”是法律术语，心理学中不常用。这次之所以使用，是因为它们具备大众所通常理解的意义，因此罗森汉把它们引入他的研究之中）。罗森汉提出，一种检验能否正确鉴别潜在病人心理健康专业能力的方法是，让一些心理健康的正常人设法进入精神卫生机构，看看那些负责诊断的人员能否发现这些人实际上是心理健康的。如果这些假病人的行为在医院里和在日常生活中一样，且医务人员没发现他们是正常的，那么这证明在精神疾病诊断中环境因素的影响大于病人自身因素。

方法

罗森汉招募了8个被试（包括他自己）来做假病人。这8个假病人（3女、5男）由1名研究生、3名心理学家、1名小儿科医生、1名精神病学家、1名画家和1名家庭主妇组成。他们的主要任务是把自己送进12所精神病院，这些精神病院位于美国东海岸和西海岸的五个州。

所有的假病人都说相同的指导语：他们打电话到医院，并预约医生。到了医院后，他们说能听到“无意义的空洞”（empty）、“空闷的”（hollow）和“砰”（thud）等声音。除了这一个症状以外，所有被试的言行完全正常，并且给问诊者的信息都是真实的（除了他们为了隐藏研究目的而改变了自己的姓名和职业）。经过面谈，所有的被试被收入不同的医院。除一人外，其余均被诊断为精神分裂症。

一旦进入医院，所有被试不再表现任何伪装症状，而且行为正常。被试不知道自己什么时候能够出院。直到他们分别说服医务人员他们心理很正常，才得以离开医院。所有被试都把自己的经历记录下来。最初，他们只是偷偷在记，但不久就发现这种隐藏是没有必要的，因为“记录行为”本身被医务人员解释为他们疾病的另外一个症状。所有被试的目标都是尽早出院，所以，他们的行为表现像是模范病人，与医务人员合作，接受所有的药物治疗（但不把药物吞下去，而是扔到马桶里冲掉）。

结果

这些假病人的住院时间从7天到52天不等，平均住院时间为19天。在此研究中关键的发现是：没有一个假病人被医务人员识破。在出院时，他们的心理状况被称为“精神分裂症缓解期”，记录在他们的病历中。他们还记录下了其他有趣的发现和结果。

尽管医院的医生、护士和其他医务人员没能鉴别出假病人，其他真病人却不那么容易被欺骗。在假病人所在的三个医院中，118个真病人中的35人对被试表示怀疑。怀疑被试不是真正的精神病人。他们这样评论：“你不是疯子！你是记者或编辑，你们是来检查医院的！”

病人（无论被试还是真病人）与医务人员之间的交流是很少的，而且常常是怪怪的。这个研究中，通过假病人做的一个试验：接近不同的医务人员，通过问一般的、正常的问题试图与他们进行言语交流（如，我什么时候可以康复？或我什么时候可以出院？）。表8—1总结了他们得到的反应。

表8—1 医生和工作人员对假病人所提问题的反应

反应	精神病学家（%）	护士和看护人员（%）
继续前行，避免正面接触	71	88
有目光接触	23	10
短时间停留和聊天	2	2
长时间停留和谈话	4	0.5

注：摘录自“On Being Sane in Insane Places, ” Science，179：255-258，1973.再印版获AAAS许可。

当假病人提问获得来自主治医师的回应时，常常有类似下面的对话形式：

假病人："劳驾一下，医生，你是否能告诉我，我什么时候可以康复？"

精神病医生："早上好，戴夫，你今天感觉如何？"

然后，医生不等待进一步的反应就继续往前走。

在医院研究中，与严重的人际交流缺乏形成鲜明对比的是，病人在药物方面从不短缺。这8个假病人在此研究中，医生共发给他们2100片药片，如前面所提到的，他们并没有吃下去。被试注意到，许多真正的病人也偷偷地把他们的药片扔到马桶里。

假病人讲到的另一个奇闻是，一个护士在满屋子全是男病人的休息室门前，解开制服的扣子，调整自己的胸罩。根据被试报告，这并不是她有意挑逗，她只是没有把精神病人当成真正的人看待。

讨论

罗森汉的研究有力地证明了，在医疗机构中，即使是受过专业训练的人员也常常不能将正常人和精神病人区分开来。根据罗森汉的研究，这是因为过于强大的精神病院环境影响了医务人员对某一个体行为的判断。一旦精神病人进入这种机构，医务人员往往会忽视病人的个体性。固有态度是："如果他们来这儿，他们就一定是疯子。"更重要的是，像罗森汉指出的"贴诊断标签"，即当一个病人被贴上"精神分裂症"的标签后，这个诊断就成为他的核心特征或人格特质。从标签被贴上并被医务人员知晓的那一瞬间，医务人员就认为病人的所有行为都源于诊断标签。因此，他们对假病人的日记缺乏关注和怀疑，只是把它当成精神病标签的另一个行为证据。

医务人员倾向于忽略病人的环境压力，而把所有的行为视为与精神病特质相关。这里有一个证明，它来自一个被试的观察。

> 一个精神病学家指着许多离吃午饭还有半小时就坐在自助餐厅外的病人，对一群年轻的住院医生说，病人的等待行为就是"精神分裂症"症状中"口唇需求"（oral-acquisitive）症状的本质特征。他似乎没有想到，在精神病院里除了吃饭，几乎没有什么事儿可做。（P.253）

除此以外，那些诊断标签甚至还歪曲了对假病人的生活史的解释。请记住，所有的被试对自己的过去和家庭都给予诚实的说明。下面是一个来源于罗森汉研究的例子，关于一个假病人对他自己历史的陈述，以及在他出院以后医生在报告中对此所做的解释。被试的真实成长史如

下：

> 童年早期，假病人与他的母亲有着密切的关系，而与父亲有些疏远。然而，从青少年时期开始，他与父亲成了亲密的朋友，而与母亲的关系变得冷淡了。他目前与妻子的关系亲密而融洽。除了偶尔争吵，冲突很少。孩子也极少挨打（P.253）。

医务人员对这一段相当正常且清白的历史的解释如下：

> 这名39岁的白人男子，在童年早期亲密关系方面显示了一个长时间的相当矛盾的历史。与母亲融洽的关系在青少年期变冷淡，与父亲关系的疏远则变得非常热切。缺乏感情的稳定性。他试图控制对妻子和孩子的情绪，但忍不住发脾气，打孩子。当他说他有几个好朋友时，让人感觉到一个相当矛盾的意识埋藏在这些关系中（P.253）。

应该指出，没有迹象表明医务人员的任何曲解是有意的，他相信诊断（此例指精神分裂症），并且对病人的生活史和行为的解释与诊断是相一致的。

研究发现的意义

罗森汉的研究震动了精神卫生专业领域，研究结果给出了两个关键因素：第一，它表明精神卫生机构不能将“心智健全”与“心智不全”区别开，正如罗森汉在他自己的文章中所讲到的：“医院本身就是一个特殊的环境，在这种环境中，行为的意义很容易被误解。结果，住院病人在这样一种环境中似乎无疑会对抗治疗”（P.257）；第二，罗森汉揭示了诊断标签的危险性。一个人一旦被贴上符合某个心理疾病的标签（如精神分裂症，抑郁症等），那个标签将掩盖他或她的所有其他特征。于是，所有行为和人格特征被看作是源于该精神障碍。这种做法最坏的影响是它可以变成自我证实，即当一个人在一段时间里被按照一定方式对待时，他或她的行为就可能开始变成被对待的那样。

罗森汉的研究除了提醒人们在诊断过程中要格外小心，也使人们认识到给病人贴标签的危险性。随着住院病人的减少，该研究强调的问题也有所减少。住院人数的下降是由于20世纪50年代抗精神病药物的发明和日益广泛应用引起的，药物的使用可有效减轻许多病人的症状，使他们能够在院外接受治疗，并生活在一个相对正常的环境里。与此并存的原因是社区精神卫生机构、危机干预中心和重视特殊问题且避免使用标签的行为疗法的发展。

但这并不意味着精神卫生专业人员已经取消了诊断标签。然而，在很大程度上，由于罗森汉的研究和其他相关研究，现在精神病标签的使用更加谨慎，而且人们对待标签时也充分考虑到它们的影响力。

问题和批评

当时，有一所研究和教学医院的医务人员在罗森汉的研究结果发表前听说了他的一些研究结果，并怀疑这些误诊出自他们医院。为了检验这一点，罗森汉通知该医院的工作人员，在接下来的三个月时间里，一个或多个假病人将会试图住进他们的精神病院。每一个医务人员都被要求用10点量表对当前的每个病人是假病人的可能性进行评定。三个月之后，该医院共收治了193人。其中41人被至少一个医务人员很自信地认为是假病人；23人被至少一个精神病专家怀疑是假病人；19人被一个精神病专家和另一个医务人员确认为假病人。罗森汉（这个狡猾的家伙）在这三个月的时间里，竟然没有派一个假病人去医院！“这个实验是有教育意义的”，罗森汉说，“它表明，当风险（在这里指声誉和诊断能力）很高的时候，将神志健全的人诊断为异常的倾向可以被反转。但存在一个不争的事实：任何一个诊断过程，本身很容易出这么多错时，它当然不是一个会让人很信赖的诊断过程。”（P.252）

罗森汉在1973年～1975年间，在12个医院重复了这项研究。每一次他都得到了相似的结果（见Greenberg，1981；Rosenhan，1975）。然而，其他研究者对罗森汉得出的结论提出了异议。斯皮策（Spitzer，1976）认为，尽管罗森汉所用的方法表明心理诊断系统是无效的，但事实并非如此。例如，通过说谎，假病人进入精神病院并不困难，因为准许入院是基于他们的口头报告（而且谁会怀疑有些人会使用欺骗的手段进入这种地方呢）。同理，你可以进入一个急救中心说自己腹痛，可能你就被诊断为得了胃肠炎、阑尾炎或胃溃疡而住院。斯皮策认为，虽然医生被欺骗了，但诊断方法并不是无效的。另外，斯皮策还指出，尽管假病人在入院后的行为正常，但是这种症状的变化在精神疾病中是常见的，也并不意味着医务人员不能诊断出欺骗行为就是无能的表现。

从1973年罗森汉的论文开始，有关心理诊断有效性的争论一直持续不断。不管还在持续的争论结果怎样，毫无疑问的是，罗森汉的研究在心理学历史上成为最有影响的心理学研究之一。

近期应用

为了说明这种持续性争论，我们将从众多应用罗森汉的研究、挑战由精神卫生专业人员制定的诊断标准有效性的研究中选取两个研究加以介绍。其中之一是托马斯·萨斯（ThomasSzasz）的研究。萨斯是一名精神病学家，在20世纪70年代早期，他已成为著名的精神病整体概念的批评家。他的观点是，精神疾病并不是病，而应该把它看成由社会和环境因素引起的“生活中的问题”。在一篇文章中，萨斯举例说明，一些被诊断为患有精神疾病的人表现出来的“疯言疯语”“并不是诊断一个人心智不全的有效证据”，这仅仅是因为一个人（精神卫生专业人员）不能理解另一个人（病人）（Szasz，1993，P.61）。

另一个研究建立在罗森汉1973年的研究基础上，这个研究探讨了在

某些现实生活情境中，人们确实会故意装出某些精神问题的症状（Broughton & Chesterman，2001）。研究中讨论的案例是关于一个被控对一名十几岁的男孩进行性侵犯的男人，当对罪犯的精神病问题进行评估时，他表现出了各式各样的精神病行为。根据进一步的检查，临床学家发现他伪装了所有的症状。研究者指出，在诊断心理障碍时，精神卫生专家传统上先假定病人陈述的正确性（就像他们对罗森汉这些假病人所做的一样）。然而，他们认为这种伪装出来的症状，“对所有的精神病学家都是一个棘手的问题，特别是外在的社会法律问题可能成为伪造精神病症状的动机时，这个问题将变得更加复杂化”（P.407）。换句话说，我们不得不留心，防止罪犯以伪造的精神病症状作为“逃避坐牢的通行证”。

当精神病人被贴上诊断标签时，他们本人会体验到怎样的感受呢？沃尔（Wahl，1999）发起了一项研究调查了1300多名精神病人，询问他们对被歧视和污名的感受。大多数被试报告受精神疾病污名影响的感受来源于多个方面，其中包括一般公众、家庭成员、教友、同事、甚至精神健康专业人员。另外，作者写道：“大多数的被试倾向于试图隐瞒他们的疾病，并非常担心其他人发现自己的精神病身份而对自己不利。他们报告说，这种经历的结果是沮丧、痛心、气愤和低自尊。他们提出，进行公众教育是减少耻辱的手段。”（P.467）

一个名为“倾听我的疯狂”的相关研究的作者（Lester & Tritter，2005）指出，一个可能有助于我们理解精神疾病患者体验的方法是，通过类比现实社会中我们所感知到的其他类型的残疾来解释他们所受的伤害。研究者认为，重度精神疾病患者与现实社会的交往通常类似于那些接受治疗的其他残障人士。通过将残障模型应用于精神疾病，患者们将有更多机会获得他们所需要的服务与帮助。

结论

精神健康方面的专家希望我们的文化能够增加对精神病的宽容和理解。尽管在许多情况下，精神病的诊断仍旧既是艺术又是科学，但可以确定的是，诊断的能力在不断提升。有可能我们永远不会去除精神病标签；它们是有效治疗心理障碍的必要部分，就像疾病的名称是诊断和治疗躯体疾病的一部分一样。然而，如果我们被标签困扰（不开玩笑），我们必须为去掉由标签所引起的污名、困窘和羞愧而继续努力。

Broughton, N., & Chesterman, P.(2001).Malingered psychosis.Journal of Forensic Psychiatry，12, 407-422.

Greenberg, J.(1981, June/July).An interview with David Rosenhan.APA Monitor, 4-5.

Lester, H.& Tritter, J.(2005) “Listen to my madness: Understanding the experiences of people with serious mental illness”.Sociology of Health & Illness, 27(5), 649–669.

Rosenhan, D., L.(1975).The contextual nature of psychiatric diagnosis.Journal of Abnormal Psychology, 84, 442-452.

Spitzer, R.L.(1976).More on pseudoscience in science and the case of the psychiatric diagnosis: A critique of D.L.Rosenhan’s “On being sane in insane places” and “The contextual nature of psychiatric diagnosis.” Archives of General Psychiatry, 33, 459-470.

Szasz, T.(1993).Crazy talk: Thought disorder or psychiatric arrogance? British Journal of Medical Psychology, 66, 61-67.

Wahl, O.(1999).Mental health consumers’ experience of stigma.Schizophrenia Bulletin, 25(3), 467-478.

研究30 你将再次获得防御

Freud, A.(1946).The ego and the mechanisms of defense.New York: International Universities Press.

在任何一部涉及改变心理学研究的历史书中，有一个重要的人物很难被忽视，那就是西格蒙德·弗洛伊德（Sigmund Freud，1856-1939）。如果没有弗洛伊德的贡献，心理学就不大可能像今天我们所知道的这样。他使我们对人类行为的解释（特别是异常行为）从中邪着魔的非理性迷信，提高到推理和科学的理性取向。因此，如果对他的工作没有一个评介，这本书就是不完整的。现在，你可能要问自己，既然弗洛伊德这么重要，为什么这篇论述是关于他的女儿—安娜·弗洛伊德（Anna Freud，1895-1982）所著的一本书？要回答这个问题，需要一点儿解释。

虽然弗洛伊德的贡献在心理学历史中是不可或缺的，对于本书而言也同样如此。但是，要把他的研究和其他人的研究一起放在这里却很困难。这是由于弗洛伊德并不是通过界定清晰的科学方法论得出他的结论的。因此，就不能像本书介绍的其他人研究那样，选择一个单独的或一系列实验代表他的工作。弗洛伊德的理论是从他对病人细致的观察，以及几十年的临床分析基础上发展起来的。因此，至少可以说，他的著作很丰富。他的英译本著作集（Freud，1953 to 1974）共有24卷！很明显，在这里能讨论的只是他著作中很少的一部分。这里所选择的文章，是弗洛伊德理论中经过时间检验的、相对没有遭到抨击的部分。在过去的一个世纪，针对弗洛伊德的理论有大量的批评。特别是近50年，从科学的角度看，他的工作受到了严重的质疑。批评者们认为，他的许多理论不能被科学所检验，或者即使能被检验，也被证明是不成立的。因此，尽管没有人怀疑弗洛伊德工作的历史重要性，但当今绝大多数的心理学家对他的人格结构理论、人格的性欲发展五阶段理论和人们心理问题的根源理论持否定态度。然而，他在某些方面的工作多年来一直得到积极的评论，目前已被相对广泛地接受。其中之一就是他的“自我防御机制”的概念。这是自我用以抵御自身产生焦虑的心理“武器”。本书收录了弗洛伊德著作中的这部分内容来代表他的工作。

西格蒙德·弗洛伊德发现的自我防御机制，是在他30多年解决心理问题的经历中逐渐发展起来的。在西格蒙德·弗洛伊德众多著作中，都没有对防御机制进行独立的系统、全面的论述。事实上，弗洛伊德把他的这项工作交由他的女儿完成，他女儿在自己的领域，特别是帮助儿童方面也是一位重要的精神分析专家。在1936年安娜的《自我和防御机制》一书第一次在德国出版之前，弗洛伊德承认了这个事实。他写道：“有大量的方法（或我们所说的机制）被自我用来执行防御功能。我的女儿——

儿童分析家写了一本与此有关的书。”（S.Freud，1936）因此，是安娜·弗洛伊德把她父亲关于防御机制的理论整合成了一本书，我们对此书进行了节选，以此展开对弗洛伊德工作的讨论。

理论假设

为了检验弗洛伊德的防御机制的概念，我们应该简单介绍一下他的人格结构理论。弗洛伊德假设，人格由三部分组成：本我、自我和超我。

在弗洛伊德的观点中，本我（id，是“it”的拉丁文简写）自出生时就有，是由人的基本生物冲动和本能组成，如饥饿、渴和性的冲动。每当这种需要得不到满足，自我都会产生强烈的信号使人找到办法满足它，并且立即这样做！本我按弗洛伊德所说的“快乐原则”行动，意味着它坚持要求即刻满足所有的欲望，不顾道德、安全、逻辑、理性等一切因素。弗洛伊德认为，存在于每个人本我中的那些黑暗的、反社会的、危险的本能冲动（特别是性冲动）会不断地寻求表现。因为本我在无意识水平运行，因此人们通常不能意识到它。然而，如果你缺乏人格中的其他部分而只有本我，弗洛伊德会预期你的行为将是不道德的、有严重的偏差或越轨，甚至对自己和别人都是致命的。

在弗洛伊德的观点中，你没有按这些危险的、越轨的方式去行动的原因，是靠你的自我和超我的发展来限制和控制本我的冲动。按照弗洛伊德的解释，自我（ego）按“现实原则”行动，即它关注现实社会和行为的结果。自我是有意识的，其工作是让本我的欲望得到满足，但是要用理性的和相对安全的方式来实现。然而，自我也受超我（superego，意思是“在自我之上”）的限制。你的超我，实质上要求自我根据已有的一套关于是非对错的准则，找到合乎道德和伦理的解决方案来满足本我的需要。弗洛伊德认为，这些道德准则是由你的父母逐渐灌输给你的，并且如果你的行为方式违反了这些准则，你的超我会用它非常有效的武器惩罚你：那就是内疚。你认出超我了吗？它就是我们通常所说的良心。弗洛伊德认为超我在意识水平和无意识两个水平上发挥作用。

弗洛伊德的人格观是一个动态的，即自我不断地试图平衡本我的需要、欲望和超我的道德要求，从而决定你的行为。接下来的这个例子说明了它是怎样工作的。想象一下，一个年轻人在一个小镇的街上漫步。时间是晚上10点钟，他正在回家的路上。突然，他意识到自己饿了，在路过一个杂货店时，他看到在一个大的橱窗里有食物，但店门已关。他的本我可能说：“看，食物！跳过橱窗拿一些。”（记住，本我想立刻满足而不顾后果）他可能没有意识到本我的建议，因为它们在意识水平之下。自我可能“听到”了这些，不过，因为它的工作是保护男孩不受危险，它可能回答：“不，那将是危险的，让我们绕到后面，破门而入，偷些食物吧！”这时，超我将表达它的愤怒：“你不能那样做，那是不道德

的，如果你那样做，我将惩罚你！”因此，这个年轻人的自我重新考虑并提了一个新建议，那是本我和超我都能接受的。“你知道，穿过四条街有一个通宵快餐店，我们去那儿买点食物吧！”假定这个男孩的心理是健康的，这就是最终反应在行为上的解决方法。

按照弗洛伊德的解释，大多数人不存在行为上的反社会或越轨行为，是因为在人格三部分之间有这个检查和平衡系统。但如果这个系统发生故障，即平衡丧失，会发生什么？可能发生的一种情况是本我的要求太强烈而自我不能充分地控制。如果不被接受的本我欲望挤进你的意识（弗洛伊德称为前意识），并开始击败你的自我，会发生什么情况？弗洛伊德认为，如果这种事情发生，你将经历一种非常不愉快的、被称为焦虑的情况，特别是他所说的游离性焦虑，虽然你感到焦虑和害怕，但你不知道为什么会有这种感觉，因为产生焦虑的原因并没有完全进入意识。

当这种焦虑状态存在时，人就不舒服，会试图改变它。要做这件事情，自我要拿起它的所谓的“自我防御机制”的“大枪”。防御机制的目的是阻止本我被禁止的冲动进入意识。如果自我成功，由本我的冲动引起的焦虑不适将缓解。防御机制通过自我欺骗和歪曲事实，使本我的欲望没有得到承认，由此避开了焦虑。

方法

弗洛伊德声称这是他多年在与自己的病人的临床交往中逐渐发现了防御机制。自从西格蒙德·弗洛伊德去世和安娜·弗洛伊德的书出版之后的这么多年来，防御机制的解释又得到了进一步完善。下文中有关防御机制的摘要是得到西格蒙德·弗洛伊德认可并又被他女儿精细化了的内容。

结果和讨论

安娜·弗洛伊德（P.44）确定了她父亲曾描述过的10种防御机制。在此将讨论今天常被使用和认可的五种原始防御机制：压抑（repression）、退行（regression）、投射（projection）、反向形成（reaction formation）和升华（sublimation）。请记住，防御机制的基本功能是改变现实以避免焦虑。

压抑

压抑是我们所使用的保护自我的最基本、最常用的防御机制。在弗洛伊德早期著作中，他把“压抑”和“防御”这两个概念互用，并且把压抑

解释为重要的、唯一的防御机制。后来，他承认压抑是许多有效防止焦虑的心理过程之一。弗洛伊德认为，个体使用压抑可以完全把烦恼冲动挤出意识层。因此，与“被禁止的”想法相联的焦虑就可避免，因为人们根本就没有意识到它们的存在。按照弗洛伊德的观点，压抑通常被用来抵御由不可接受的性欲望产生的焦虑。例如，一个妇女对他的父亲有性欲的感觉，如果她意识到这种冲动，她将表现出强烈的焦虑。为了避免焦虑，她可能压抑她的不可接受的性欲望，迫使它们全部进入无意识。这并不意味着她的欲望消失了，只是因为它们被压抑而不能产生焦虑。

你可能觉得奇怪，如果这些想法存在于无意识之中，又怎么会被发现？根据弗洛伊德的解释，这些隐藏的冲突可以通过口误、梦、自由联想或催眠术等精神分析的各种技术（如自由联想或催眠术）被揭示出来。此外，压抑的欲望可产生心理问题，并以神经症的形式表现出来。例如，上面提到的那个妇女压抑了对父亲的性欲望。她可能以另外的形式表达了这种冲动，即她无意识中试图解除对父亲的情感冲突，而不断地陷入与男性的失败关系。

退行

退行是一种被自我用来防止焦虑的防御机制。这种防御使人退回到发展的早期阶段中那种要求很少和相对安全的行为。通常，当第二个孩子来到一个家庭，大孩子将退行，用小孩子的说话方式，想要奶瓶甚至尿床。成人也可能用退行机制。想一想，一个男人感受到了“中年危机”（midlife crisis），他害怕变老和死去，为了避免与这些无意识恐惧相联的焦虑，他可能退行到青年时期，变得不负责任，开赛车兜风，试着与年轻女性约会，甚至是吃儿童的食品。另一个退行的例子是，每当婚姻中出现问题时，已婚的成人就会搬回家跟妈妈住。

投射

想象一下你的自我正被你的本我挑战。虽然你不知道为什么，但你感到非常焦虑。如果你的自我用投射机制消除焦虑，你将从别人的行为中看到你无意识的欲望，即你把你的冲动投射到别人身上。理论上，这样可以把引起焦虑的感觉外化，进而减少焦虑。你将不会意识到你做的这些，而被你投射的人也不可能因你的谴责而有罪恶感。安娜举的一个例子是：一位丈夫有了对他妻子不忠的冲动（P.120），他可能甚至没有意识到这种欲望，但这种欲望从他的本我不知不觉地爬上来，并引起焦虑。为了防止焦虑，丈夫把他的欲望投射给他妻子，变得善妒，谴责她有外遇，即使没有证据支持他的谴责。另一个例子是，一位妇女害怕变老，开始指出她的朋友和熟人看起来多么老。这些例子中的个体并不是在表演或说谎，而是真实地相信他们的投射。如果他们不这样想，对焦虑的防御将会失败。

反向形成

弗洛伊德用莎士比亚的《哈姆雷特》中的一句话来举例说明反向形成的防御机制。当哈姆雷特的妈妈看完演出的一幕，对哈姆雷特说："我认为，这个女士申辩得太多了"。当一个人感受到不可接受的、无意识中的"邪恶"冲动，为了避免这些所引起的焦虑，人们表现出与本我真实欲望相反的行为。安娜指出，这些行为通常是夸张的，甚至是过分的。人们采取表面的、完全没有本我真实欲望的态度和行为来阻止焦虑。除非本我和自我的冲突以某种方式解决，否则，反向形成倾向于成为个体人格的永久的部分。举这样一个例子，再想一想刚才那个丈夫，他潜意识里想得到别的女人，如果他是用反向形成而不是用投射来防止焦虑，那么他会变得过度殷勤地对待他的妻子，并送她礼物，宣布他忠贞不移的爱情。另一个例子来自于被称为"怒殴同性恋"（gay bashing）暴力犯罪的令人不安的新闻报道。按照弗洛伊德的解释，一个体验到无意识的同性恋欲望（这种欲望是他所恐惧的，因为社会不认同同性恋取向）的男人，可能会有这种攻击或打击同性恋的极端反向的行为，以掩藏他真实的欲望和与此相联的焦虑（这个概念在后面部分中会进一步讨论）。

升华

西格蒙德·弗洛伊德和安娜·弗洛伊德两人都把大多数防御机制，包括之前描述的四种，作为心理调适有问题（神经症）的指标。相反，他们不但把升华这种防御机制看作是正常的，而且是值得利用的。当人们使用升华机制，他们就找到了以社会认可的方式来转化无意识中被禁止的欲望所产生的焦虑的能量。西格蒙德·弗洛伊德主张既然每个人的本我都包含这些欲望，升华就是多彩和健康生活所必需的一部分。另外他相信，那些最强烈的欲望可用不同的方式升华。某个人有强烈的攻击冲动，可以通过参加对抗运动或成为外科医生来得到升华；一个小女孩喜欢骑马，可以解释为不可接受的性欲望的升华；一个好带着色情的目光注视人体的男人，可以通过成为一名裸体像画家或雕塑家来升华他的感情。

弗洛伊德提出，所有被我们称之为"文明"的现象都可能是通过升华机制而来的。按照他的观点，人类可以升华他们原始的、生物性的欲望和冲动，引导他们转向建设文明的社会。然而，弗洛伊德认为，有时人类的无意识的力量会击败我们的"集体自我"（collective ego），而这些原始的、兽性的欲望就会以不文明的、野蛮的形式爆发出来，例如战争。但是，不管怎样，也只有通过升华，文明才能存在（S.Freud，1936）。

意义及近期应用

虽然安娜·弗洛伊德在她的书中清楚地写道，防御机制的使用常常与神经症行为并存，但情况并不总是如此。几乎每个人在他的生活中都会不时地应用各种防御机制，尤其是帮助他们应对压力增加的时期。防御机制帮助我们减少焦虑，保持一个积极的自我意象。某些防御机制的使用甚至被证明可以减少非健康的心理活动。例如，一项研究发现投射防御机制的使用与血压的降低相关（Cramer，2003）。不过，防御机制包含了自我欺骗和歪曲事实，如果过度使用，会产生消极的结果。例如，一个人每次都用退行机制，生活中的难题就会越积越多，他可能永远都不会自己采用必要的策略去处理问题和解决问题。结果，他作为一个整体人的发展会被抑制。而且，弗洛伊德和许多其他心理学家主张，当某些特定冲突引起的焦虑被压抑后，它就会在某些时候以其他形式表现出来，如恐怖症、惊恐发作或强迫症。

当下，绝大多数研究者对弗洛伊德大部分的理论都有所质疑，包括他的自我防御机制的概念。防御机制真的存在吗？真的能通过阻止本我的冲动来阻止焦虑的产生吗？也许最常引证的对于弗洛伊德所有工作的批评是：说得好听一点是很难科学地验证它；通常的批评则是不可能科学验证它。许多研究试图清楚地证明弗洛伊德的各种概念的存在。研究结果是多种多样的，其中一些观点已经得到一些科学的验证（见Cramer，2007），而另一些被证伪，还有一些无法进行科学研究（见Fisher & Greenberg，1977，1995）。

有一项吸引人的研究所得到的支持性证据来自于同性恋恐怖症（homophobia），即某些异性恋者所表现的对男同性恋和女同性恋的不合理的恐惧、回避和偏见，其实可以被看作是一种反向形成，即为了避免由自我压抑的同性恋倾向引起极端的焦虑所采用的防御机制（Adams，Wright & Lohr，1996）。在这个研究中，一组男性接受一种评估同性恋恐怖症水平的纸笔测验，然后被分成两组：同性恋恐怖症组和非同性恋恐怖症组。接着，所有的被试观看描述异性恋、男同性恋或女同性恋表演性活动的录像，当他们观看这些录像的时候，同时监测他们性唤起的生理信号。两组被试唯一的不同是，他们观看男同性恋的录像时的生理情况。“结果显示，同性恋恐怖症的男性在性唤起方面表现明显增加，而非同性恋恐怖症的男性没有这些表现。”（P.443）事实上，当观看男性同性恋的录像时，非同性恋恐怖症组中66%的被试没有明显的性唤起信号，而同性恋恐怖症组仅有20%的被试表现出很低的性唤起，或没有明显的唤起信号。而且，当要求同性恋恐怖症组的男性主观评价自己的性唤起水平时，他们低估了自己在对同性恋录像中的性唤起的水平。这项研究结果与安娜·弗洛伊德对反向形成的描述相吻合。同时，也为暴力反对同性恋的可能解释提供了支持。

结论

从本文中的研究中可以看出，心理学家对防御机制的科学兴趣在各

分支学科不断增加，包括认知心理学、人类发展心理学、人格心理学和社会心理学（参见Cramer，2000）。通过对防御机制的认识和了解，你对人们行为原因的洞察能力将明显提高。如果你在自己的头脑中保存一份防御机制的条目表，你就可以开始在别人或自己身上发现它们。顺便说一下，如果你认为某人正在使用防御机制，记住，他或她这样做的原因是为了避免令人讨厌的焦虑。所以，提醒他们注意到这一点可能并不是一个好主意。虽然在你与别人的交往中，防御机制的知识是一个强有力的工具，但必须小心、谨慎地使用这些知识。

通过详细查看近期关于心理分析理论的主要学术著作，你会很容易地体验到安娜·弗洛伊德对她父亲的那些防御机制概念的综合、分析对你的持久影响。的确，你所看到的大多数的弗洛伊德的引证都与西格蒙德有关。但是，当讨论转向防御机制时，安娜·弗洛伊德在1946年出版的书及其各种修订本被认为是论述此话题的权威性著作。

Adams, H., Wright, L., & Lohr, B.(1996).Is homophobia associated with homosexual arousal? Journal of Abnormal Psychology，105(3), 440–445.

Cramer, P.(2003).Defense mechanisms and physiological reactivity to stress.Journal of Personality, 71, 221–244.

Cramer, P.(2007).Protecting the self: Defense mechanisms in action.New York: Guilford Press.

Fisher, S., & Greenberg, R.(1977).The scientific credibility of Freud’s theories and therapy.New York: Basic Books.

Fisher, S., & Greenberg, R.(1995).Freud scientifically reappraised: Testing the theories and therapy.NewYork: Wiley.

Freud, S.(1936).A disturbance of memory on the Acropolis.London: Hogarth Press.

Freud, S.(1953 to 1974).The Standard Edition of the Complete Psychological Works of Sigmund Freud.London: Hogarth Press.

研究31 习得性抑郁

Seligman, M.E.P., & Maier, S.F.(1967).Failure to escape traumatic shock.Journal of Experimental Psychology, 74，1-9

如果你和大多数人一样，你就会期望你的行为产生特定的结果。你的预期既会使你的行为朝着令人满意的结果迈进，又会尽量避免令人失望的结果。换句话说，你的行为至少部分由一种信念所决定，即你相信自己的行为会产生某种后果，而这种后果与行为有对应关系。

让我们设想一下，假如你不喜欢现在的工作，于是你开始着手准备换个工作。你与周围的熟人联系，翻阅你感兴趣的招聘广告，为获取新的技能参加业余培训，等等。所有这些行为都由你的信念所引发，即你的努力最终会给你带来好工作和幸福的生活。人际关系也是如此，如果你正处于一种不良的人际关系中，这种人际关系带有侮辱性或使你不愉快，你将希望采取必要的行动改变它或终止它，因为你期望成功地建立一种令人满意的人际关系。

所有这些都与能力和控制力有关。许多人相信自己有足够的能力，并能控制发生在他们身上的一切，至少在有些时候是这样，因为他们在过去曾对某些情况进行过控制并取得了成功。他们相信能够通过自己的努力而达到目标。如果一个人觉得自己缺乏能力和控制力，那么剩下的就只有无助和绝望了。如果你感到自己被一份不满意的工作所困扰，却又找不到另一份工作，或不能学习新技术来提高自己的职业水平，那么你将不太可能为改变工作而作出必要的努力。如果你太依赖某人，而你与他的关系已经破裂，又觉得无力去修复或结束它，那么你便只能无奈地处于这种关系中并忍受煎熬。

能力与控制的知觉对身心健康是非常重要的（参见研究20中兰格和罗丁关于养老院中老年人自制力的研究讨论）。想象一下，如果你突然发现自己不再有使生活发生改变的能力和控制力，在你身上发生的事情与你的行为无关，你将有何感受？你可能将感到无助和无望，并且放弃一切尝试。换句话说，你将变得抑郁。

马丁·塞里格曼（Martin Seligman）是非常著名和有影响力的行为心理学家。他认为，我们对能力和控制的知觉是从经验中习得的。他相信，当一个人控制特定事件的努力遭受多次失败后，他（或她）将停止这种尝试。如果这种情形出现得太过频繁，这个人就会把这种控制缺失的知觉泛化到所有的情景中，甚至泛化到实际上能控制的情况下。于是，他（或她）开始感到自己像一颗“命运的棋子”（pawn of fate）任人摆布，无助而抑郁，塞里格曼把这种抑郁的产生原因称为“习得性无

助”（learned helplessness）。塞里格曼在宾夕法尼亚大学以狗为被试，通过一系列现在被视为经典的实验，发展了他的理论。我们下面将要讨论的这项研究由马丁·塞里格曼与史蒂文·梅尔（Steven Maier）共同完成，它被认为是对其理论有决定性作用的最早论据。

理论假设

塞里格曼从一项有关学习的早期实验中获悉，狗在受到既不能控制也无法逃脱的电击一段时间后，即便逃离的机会唾手可得，也学不会逃走。你想一想，这在一个行为主义者看来是多么奇怪的事情。在实验室中，对狗的惩罚是电击，但这种电击不会对其构成伤害。然后，把狗放在一个“梭箱”（shuttle box）里，那是一个大箱子，由一块隔板分割为两部分。在箱子一边的地板上通电。狗感到箱子的一边有电流时，便只需越过隔板，跳到箱子的另一边即可避开电击。通常，狗和其他动物都能很快学会这种逃脱行为（不难看出其中的原因）。事实上，如果有一个信号（如一个闪光灯或一个蜂鸣器）警告狗电流即将来临，狗将学会在电击前跳过隔板而完全避开它。然而，在塞里格曼的实验中，在梭箱里的狗在经历无法逃脱的电击后，它们就学不会这种“逃脱—回避”（escape–avoidance）的行为了。

塞里格曼的理论认为，动物在学习控制不愉快刺激的过程中存在着某些因素，这些因素决定了它们后来的学习。换句话说，这些狗在先前的电击经历中已经懂得自己的行为不能改变电击结果。因此，当它们处于新的环境中时，即便它们有能力逃脱——做出控制行为——它们也会放弃。它们已习得了无助感。

为了检验该理论，塞里格曼和梅尔计划研究可控电击与不可控电击如何对后来学习回避电击产生影响。

方法

该研究是本书中几个用动物作被试的经典研究之一。而该研究所引起的关于动物研究的伦理问题，可能比其他几项研究都要多。为了检验一种心理学理论，就要让狗忍受令其痛苦的电击（虽然没有身体上的伤害）。这种做法在伦理上是否恰当，是每个研究者和心理学学习者必须面对的问题（这个问题我们将在讨论塞里格曼的研究结果后再进行论述）。

实验的被试是24只“杂交犬，它们肩高15~19英寸，体重25~29英磅”（P.2）。它们被分为3组，每组8只。一组是“可逃脱组”，另一组是“不可逃脱组”，第三组是“无束缚的控制组”。

可逃脱组的狗和不可逃脱组的狗均被单独安置并套上狗套，这种套子与巴甫洛夫设计的实验装置相似（参见研究9中对巴甫洛夫实验方法的论述）；虽然狗受到约束，但并不是完全不能移动。在狗头部的两边各有一个鞍垫（panel），以保持头部面朝正前方。狗可移动头部，以挤压两边的鞍垫。可逃脱组的狗受到电击后，它可以通过挤压头部两边的鞍垫终止电击。不可逃脱组的狗与可逃脱组的狗一一配对（这是一种称为“匹配”的实验程序），然后在同一时间给每一对狗施加完全相同的电击，但不可逃脱组的狗不能控制电击。无论这些狗做什么，电击都将持续，直到可逃脱组的狗挤压鞍垫终止电击为止。这样就能确保两组狗接受电击的时间和强度完全相同，其唯一不同的是一组狗有能力终止电击，而另一组却不能。8只控制组的狗在实验的这一阶段不接受任何电击。

可逃脱组的狗和不可逃脱组的狗在90秒的时间里均接受了64次电击。可逃脱组的狗很快学会了挤压旁边的鞍垫来终止电击（既为它们自己，又为不可逃脱组）。24小时以后，所有的狗被放入前面所述的梭箱中，箱子两边都安装有灯，当箱子一边的灯光熄灭时，电流将在10秒种后通过箱子的底部。如果狗在10秒内跳过隔板，它就能完全避免电击。如果不这样做，它将持续遭受电击，直到它跳过隔板，或直到60秒钟电击结束。每只狗在此梭箱中进行10次试验。

研究者根据以下指标对学习程度进行了测量：（a）从灯光熄灭到狗跳过隔板平均需要多长时间；（b）完全没有学会逃脱电击的狗在每组中所占的比率。另外，7天后不可逃脱组的狗在梭箱中再次接受10次额外测试，以评价该实验处理的持续效果。

结果

在64次电击的过程中，可逃脱组的狗用于挤压鞍垫并停止电击的时间迅速缩短；而不可逃脱组的挤压鞍垫行为在30次尝试后便完全停止。

图8—1表示了在梭箱中进行的全部实验中，三组狗逃脱所用的平均时间。记住，这个时间是指灯光熄灭到狗跳过隔板之间的时间。可以看出，不可逃脱组的狗与其他两组间的狗存在显著差异；但可逃脱组的狗与控制组的狗之间的差异无显著性。图8—2表示10次尝试中至少9次不能跳过隔板并避免电击的狗在每组中所占的比率。可逃脱组的狗与不可逃脱组的狗之间也存在非常显著的差异。不可逃脱组中有6只狗在9次甚至全部10次实验中完全失败。7天后，这6只狗被放入梭箱中再次进行测试；结果，6只狗中的5只没能在任何一次试验中逃脱电击。

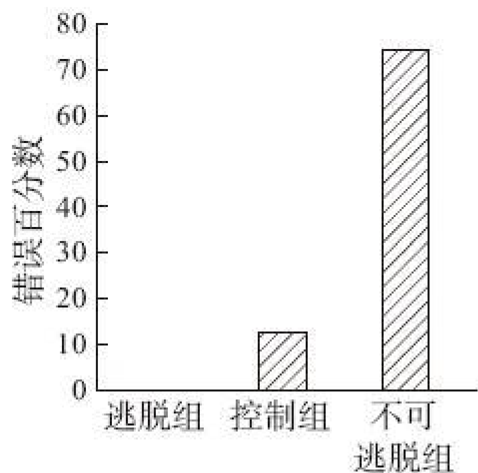


图8—1 在梭箱中成功逃避所需的

（引自P.3）

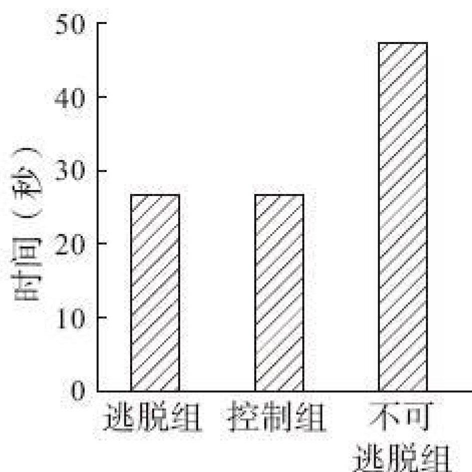


图8—2 在梭箱中被试未成功习得逃避电击的百分比

（引自P.3）

讨论

因为可逃脱组的狗与不可逃脱组的狗之间唯一的不同是：狗能否主动终止电击。因此，塞里格曼和梅尔得出结论认为，一定是这种控制因素导致了两组狗在梭箱中学习逃脱电击时表现出明显不同。换句话说，可逃脱组的狗在梭箱中能正常学会新技能的原因是，它们在前一阶段已

习得自己的行为与电击终止之间存在相关。因此，它们能主动地跳过隔板并逃脱电击。而不可逃脱组的狗在前一阶段的行为与电击的终止毫无干系。因此，在梭箱中它们并不认为自己的行为能终止电击，故不会主动尝试逃脱。正如塞里格曼和梅尔所预言的，它们习得了无助感。

偶然地，不可逃脱组的一只狗在梭箱中做了一次成功的逃脱。然而，在下一次试验中，它又恢复到无助状态。塞里格曼和梅尔对此的解释是，即便是在一次成功的经历之后，前一阶段的无效行为依旧阻止它们在新的情境（梭箱）中形成终止电击的新行为方式（跳过隔板）。

塞里格曼和梅尔在后继的实验中报告了另一些有趣的发现。在第二项研究中，首先将狗放置在可逃脱的情境下，在此情境下，狗可挤压鞍垫以终止电击，然后将其转移到不可逃脱的条件下，最后将狗放入梭箱中接受10次实验。处于不可逃脱条件下时，狗不断尝试挤压鞍垫，并没有像在第一项研究中的狗那么快地放弃了尝试。而且，它们在梭箱中都成功地学会了逃脱和回避电击。这表明，一旦动物习得了有效的行为，随后的失败经历不足以消除它们改变自己命运的动机。

后续研究

当然，塞里格曼也想做你可能已经在脑子里进行的事情：把这些研究结果应用于人类。在随后的研究中，他声称人类的抑郁发展与动物习得性无助的形成过程非常相似。在这两种情形下，他们（它们）都表现出被动、消极、坐以待毙、缺乏进取心、学习某些成功行为极为缓慢、体重减少和社会性退缩等行为。无助的狗和抑郁的人都认为从以往的特殊经历中习得自己的行为是徒劳的。无论狗做什么，它都无法逃脱电击；而人也有无法控制的事件，如爱人的去世、父母的粗暴、失业或严重的疾病（Seligman，1975）。

习得性无助导致的人类抑郁可以产生比抑郁本身更严重的后果。研究表明，许多老年人，像生活在养老院中的老人，由于各种原因而失去了对自己日常生活的控制力，他们的健康状况比那些能保持这种控制力的老人差，死亡的概率也更大。另外，一些研究也表明，不受人力控制的应激事件在诸如癌症这样的严重疾病的形成中起重要的作用。一项研究发现，在过去的几年中遭受丧偶、失业、失去威望等事件都将增加一个人的癌症患病概率（Horn & Picard，1979）。在医院里，医生和其他医务人员都希望病人能心平气和地配合医生，心甘情愿地把自己的命运交到这些医学权威的手中。病人认为他们想要尽快康复，就必须完全听从医生和护士的指导。一位著名的健康心理学家曾指出，做一名“好的住院病人”意味着病人必然是被动的，而且必须放弃所有控制力。这其实是为病人创造了一种习得性无助的境况，此后即使控制力有可能对继续康复起作用，这些病人也已丧失了使用控制力的能力（Taylor，1979）。

由芬克尔斯坦和雷米（Finkelstein & Ramey，1977）完成的一项引人注目的研究可作为习得性无助的进一步证据。研究者在所有婴儿床的上方架设旋转的可动装置（rotating mobiles）。有一组婴儿，其头枕着对压力敏感的（pressure-sensitive）特制枕头，以便他们可以通过转动头部来控制可动装置的旋转。另一组婴儿，其床的上方虽然有同样的可动装置，但它们被设定为随机旋转，不受婴儿的任何控制。可控制枕头组的婴儿每天有10分钟时间接触这种可动装置。经过两周，这组婴儿掌握了非常熟练的技术，通过转动自己的头使可动装置旋转。然而，更重要的发现接踵而来——后来实验者把不可控制枕头组的枕头全换成可控制的，并给予这组婴儿比第一组更多的学习时间，他们却全都无法学会控制可动装置。第一种情境下的经历已使这些婴儿了解到自己的行为是无效的，并且这种认识迁移到可控制的新情境下。对于可动装置，婴儿们已习得了无助。

近期应用

塞里格曼关于习得性无助的研究一直影响着当代的研究，并在许多领域引起争论。其他一些研究者致力于增加我们对个人控制力在生活事件中的重要性的认识，塞里格曼的观点与这些研究者的相吻合。

关于这种广泛影响的一个非常糟糕的例子是恐怖袭击引发的大面积恐慌和之后的“反恐战争”。2001年9月11日发生了对美国世贸中心和五角大楼的恐怖袭击。这些恐怖事件产生的心理上的影响在美国和全世界范围内蔓延。这些症状包括持续增长的焦虑、愤怒、紧张、酒精滥用、对外部事件控制感丧失以及无助感（CDC，2002）。事实上，恐怖分子的袭击目的之一就是让人们产生脆弱和无助感。一位临床心理学家对袭击后的影响作出如下总结：

> 恐怖主义威胁为焦虑和抑郁的出现营造了标准的心理氛围。心理学家称之为“预期焦虑”，就像谚语中描述的那样，一直在等待另外一只鞋掉下来。现在这个情况下即是等待恐怖分子的炸弹爆炸。加上“习得性无助”这个成分，即面对阻止恐怖主义，你意识到自己能做的很少或者什么也做不了，会使抑郁、易受伤害、控制感缺失进一步加剧。这些正好是自“9·11”袭击事件后，我们大家所感受到的状态。这些就是“新常态”和“9·11综合征”的特征（Braikery2002）。

有趣的是，近期的一个研究表明，间接地体验到创伤性事件，经过一段时间后，可能会有一些心理获益（Swickert等，2006）。尽管作者没有否认或轻视“9·11”事件目击者所经受的极度痛苦所带来的心理影响，但他们指出在一些个体身上存在着相反的结果，这些人被称为创伤后成长。研究者指出，以往的研究假定“当人们对自己、他人和将来的基本假设受到挑战时，会出现创伤后成长。作为对这个挑战的回应，心理受创伤的个体可能会试图从他们的经历中寻找意义。因此，有些人常会发现他们从创伤事件中获益”（P.566）。你可能会问，这种经历会产生哪些可能的益处？作者指出，另一个研究已经给出答案。在经

历“9·11”事件后，个体的一系列积极特征都有所增强，其中包括感激、希望、仁慈、领导力、爱、灵性和团队合作。研究者说，那些间接目击了袭击的人在事件发生不久后报告了相似的获益，但这些影响会随着时间的流逝而减弱。

结论

现在，我们再回到实验的伦理问题上。对我们中的大多数人而言，读到在心理学实验室中用动物，特别是把狗或其他常成为我们宠物的动物作为被试进行电击实验是很难接受的。在过去的许多年中，实验的伦理标准已得到发展，以确保实验者更人道地对待动物（参见本书前言中就该问题的论述）。然而，在科学领域内外，还有许多人认为这些标准是不够的。一些人提倡在心理学界、医学界和所有的科学界彻底取缔以动物为被试的研究。在此问题上，无论你的个人立场如何，你应该思考的问题是，这些研究的结果是否扩展了我们的知识，减轻了人类的痛苦，并提高了人们的生活质量，所有这些收益是否足以为其所用方法的合理性进行充分辩护？

就塞里格曼和梅尔的这项研究，请你自己思考一下以上这些问题。他们的研究创立了解释为什么有些人变得无助、无望和抑郁的理论。塞里格曼随后继续发展了一种被人们广泛接受的治疗抑郁的模式和方法。他的理论经过多年的不断完善和补充，现已能对在明确界定的条件下发生的抑郁进行更精确的治疗，比如爱人的死亡、重大的自然和人为灾难。

例如，通过塞里格曼的研究，我们现在了解到，如果个体学会把自己的控制力缺失归因于：（a）永久性的，而不是暂时性的；（b）自己的内在人格因素（而不是情境因素）；（c）渗透到他们生活中的许多方面（参见Abramson，Seligman, & Teasdale，1978），那么，个体最有可能变得抑郁。认识到这些后，治疗师和咨询师就能更好地理解、干预和治疗严重的抑郁症病人了。

这些知识能否表明在习得性无助的早期研究中采用的方法是合理的呢？这个棘手的问题就由你自己来决定吧！

Abramson, L., Seligman, M., & Teasdale, J.(1978).Learned helplessness in humans: Critique andreformulation.Journal of Abnormal Psychology, 87, 49–74.

Braiker, H.(2002).“The September 11 syndrome”—A nation still on edge.Retrieved September15, 2003, from http://www.harrietbraiker.com/OpEd.htm

Centers for Disease Control (CDC).(2002).Psychological and emotional effects of the September 11 attacks on the World Trade Center—Connecticut, New Jersey, and New York, 2001.Centers for Disease Control and Prevention: Morbidity and Mortality Weekly Report, 51, 784–786.

Finkelstein, N., & Ramey, C.(1977).Learning to control the environment in infancy.Child Development, 48, 806–819.

Horn, R., & Picard, R.(1979).Psychosocial risk factors for lung cancer.Psychosomatic Medicine, 41, 503–514.

Seligman, M.(1975).Helplessness: On depression, development, and death.San Francisco, CA: Freeman.

Swickert, R., Hittner, J., DeRoma, V., & Saylor, C. (2006).Responses to the September 11, 2001, terrorist attacks: Experience of an indirect traumatic event and its relationship with perceived benefits.The Journal of Psychology ，140(6), 565–577.

Taylor, S.(1979).Hospital patient behavior: Reactance, helplessness, or control? Journal of SocialIssues, 35 ，156–184.

研究32 拥挤导致行为失常

Calhoun, J.B.(1962).Population density and social pathology.Scientific American, 206，139-148.

近几十年来，拥挤对行为的影响始终令心理学家很感兴趣。可能你也注意到，当自己身处拥挤的人群中时，你的情绪和行为会有所变化。你变得孤单离群、不引人注目，或是想法儿抽身离开，抑或是变得易激惹和好攻击。

本章中这篇研究报告在题目中用了“人口密度”（population density）一词，而没有用“拥挤”（crowding）。这两个词看起来很相似，但心理学家对两者有明确的区分。人口密度被定义为一定空间内个体的数量。如果一间12×12英尺的房间内有20个人，这房间就被认为是人口密度很高。而拥挤则是由不同程度的密度造成的个体主观体验。如果你试图专心于一项困难的任务，而却与20个人待在这样一间屋子里，这时你可能感到太拥挤了；相反，如果你在上述同样的房间里与20位朋友聚会，你可能一点也不感到拥挤。

行为学家研究密度和拥挤效应的一种方法是，观察那些已经很拥挤的地方，如曼哈顿、墨西哥城、某些住宅区和监狱等。用这种方法进行研究的问题是，在这些地方除了人口密度外，还有许多其他因素也会影响行为。例如，如果我们发现在一个拥挤的贫民窟犯罪率很高，我们无法断言拥挤就是犯罪发生的原因，可能是贫穷，可能是吸毒率较高，也可能是所有这些及其他因素与拥挤因素共同导致了高犯罪率。

另一种研究拥挤的方法是让人类被试在高密度条件下待较短的一段时间，并研究他们的行为反应。如果是让被试长时间处于该状况则是不符合伦理的。尽管采用这种方法，研究者可以控制更多因素，并分离出拥挤这一因素对行为产生的影响，但这与现实中拥挤的生活环境相比，它还不是很真实，因为现实中的拥挤往往持续很长一段时间。然而，这两种研究方法都得到了关于拥挤的一些有趣的发现，我们稍后将在本文中对此进行讨论。

只为研究的需要而让人长时间处于拥挤环境中，这在伦理学上是不可行的（因为存在应激和其他潜在的损伤后果），因此研究者们使用了第三种方法来研究密度的影响：用动物作被试进行研究（参见前言部分关于动物研究的论述）。这类研究中最早最关键的系列研究之一是由约翰·B·卡尔霍恩（John B Calhoun，1917~1995）在20世纪60年代早期年进行的。卡尔霍恩让几组白鼠的数量增加到在狭小空间中正常数量的2倍（靠它们自己繁殖），然后观察它们在这一环境中存活16个月所表

现出来的“社会”行为。

理论假设

卡尔霍恩特别想了解高人口密度对社会行为的影响。虽然把白鼠看作社会性动物对你来说似乎有点奇怪，但它们在自然环境中会以许多社会化的方式来互动。

在探讨卡尔霍恩的这项研究以前，有必要回顾他更早几年所做的一项研究。卡尔霍恩把一群白鼠关在一个1/4英亩的安全而封闭的户外空间内，给予它们充足的食物，理想的、受保护的筑巢空地，没有天敌，所有的疾病被控制在最低限度。换句话说，这是一个老鼠的天堂。卡尔霍恩早期的研究重点仅仅是研究在一个不存在控制过度繁殖的自然因素（例如掠夺者、疾病等）的环境中，白鼠的繁殖率为多少。27个月后，仅剩150只成年鼠。这很让人吃惊，因为在这样理想的环境中，考虑到成年鼠的低死亡率和通常的繁殖率，在这一段时间的累积下，卡尔霍恩应该看到5000只成年鼠才对。卡尔霍恩认识到，白鼠数量有限的原因在于幼鼠极高的死亡率。显然，在这150只白鼠中，繁殖行为和母性行为由于社会相互影响的压力而严重改变了，因此几乎没有幼鼠可活到成年。虽然150只白鼠生活在1/4英亩的区域内，看起来似乎不是特别拥挤，但这种拥挤程度已足以导致明显的行为变化。

这一发现促使卡尔霍恩在实验室中设计控制更严格、更易观察的实验条件，以便进一步研究在这样的高密度条件下，白鼠会发生什么样的行为变化。换句话说，他已经观察到了现象，而现在想弄清原因。

方法

在由三项研究组成的系列研究中，研究者将成年白鼠放入10×14英尺的实验空间里，实验空间被分成四个部分或围圈（pen）（见图8—3），它们之间有通道可让白鼠从圈1到圈2，从圈2到圈3，从圈3再到圈4，但圈1和圈4之间无法直接相通。因此，圈1和圈4就是“末端围圈”。如果一只白鼠想从圈1到圈4中去，它就不得不经过圈2和圈3，每部分的隔板是通电的，因此，白鼠很快便知道，它们不能爬过隔板。

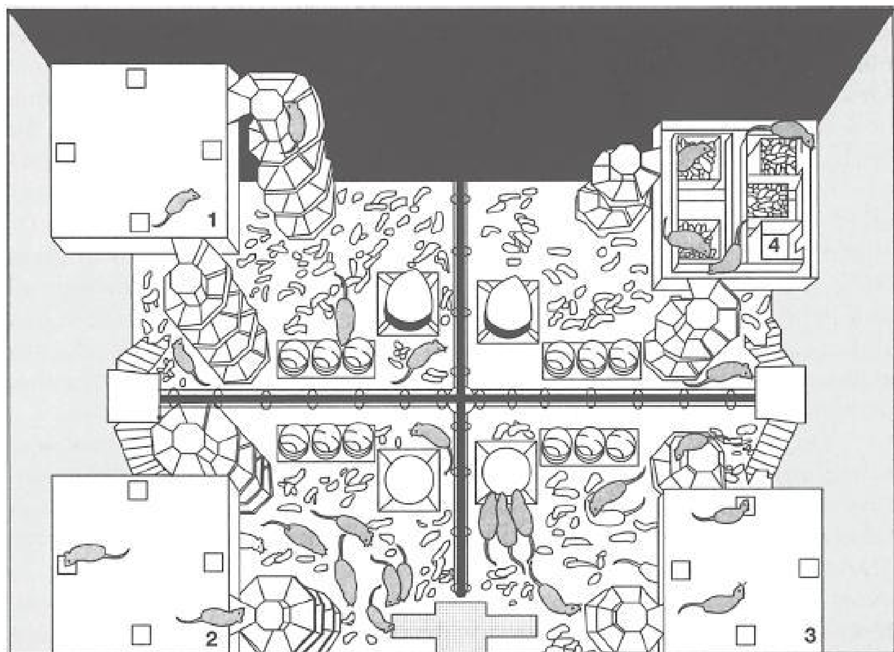


图8—3 卡尔霍恩的拥挤研究中实验房间的安排示意图

这些围圈里有食物、水和供白鼠筑巢用的围篱。白鼠有足够的食物、水和筑巢用的材料。在房间的天花板上有一个观测窗，用于研究小组观察和记录白鼠的行为。

通过几年对白鼠的研究，卡尔霍恩知道，在正常情况下，这个特殊的白鼠种群是12只成年白鼠聚居在一起的。因此，观察室的大小是每个围圈能够容纳12只白鼠，共48只。白鼠被放进观察室后，允许自由繁殖，直至接近正常密度的2倍，达到80只。一旦白鼠的数量达到80只时，多余的幼年白鼠在断奶后将被转移，以使白鼠的数量保持恒定。

有了这种实验装置，剩下的所有工作就是要对这些处于拥挤环境中的动物观察一段时间，并记录他们的行为，这些观察持续了16个月。

结果

对于白鼠来说，实验中的种群密度水平并不极端。事实上，这种密度只是中等程度。如果还想增加白鼠的话，每个围圈大约还可增加20只，但事实并非如此。当雄白鼠成熟以后，它们像在自然环境中一样，为了社会地位而开始相互争斗。这种争斗发生在所有的围圈中，但每个围圈中的结果并不完全相同，因为两个末端围圈（1、4）只有一个进出口。因此，当这两个围圈中的一只雄性白鼠赢得了战争，处于统治地位时，它将捍卫它的地位和领土（整个围圈），其方法很简单，即守住唯

一的那个入口，攻击任何其他胆敢冒险通过那个通道的雄鼠。于是，其他的雄鼠都被逐出，每个末端围圈中最终只剩下一只雄鼠，当然，那里并不是只有它自己，因为在四个围圈中雌鼠的分布较为平均。因此，围圈1和4的“统治者”们每只霸占8到12只雌鼠。而且这些雄鼠没有心存侥幸。为了防止外敌侵入，雄鼠直接睡在通道的入口处，并始终处于警戒状态。

偶尔，有些别的雄鼠也会进入末端围圈，但它们都处于从属地位，大部分时间都与雌鼠待在窝里睡觉，出去仅是为了觅食。它们不想与雌鼠交配，在这些围圈中的雌鼠基本上起着母亲的作用，它们筑建舒适的窝、养育和保护后代。换句话说，在末端围圈中生活的大部分白鼠相对正常，繁殖行为也很成功，幼鼠中有一半可存活到成年。

其余约60只左右的白鼠挤在中间的两个围圈中。因为这两个围圈都有中央供食和供水装置，因此，这些白鼠彼此间有很多相互接触的机会。在围圈2和3中的白鼠有一种行为现象，卡尔霍恩称之为“行为沉沦”（behavioral sink）。行为沉沦是“把超乎寻常数量的动物聚集在一起所产生的任何行为变化过程的结果。这个术语暗示着它本身对健康不利，这并非偶然，行为沉沦的确会使在一个集体内发生的所有形式的病理症状恶化”（P.144）。让我们看一看他所观察到的一些极端的、病理性的行为。

1.攻击

在野生环境中，正常雄鼠会为了维护自己在社会阶层中的统治地位与其他雄鼠进行争斗。这种争斗在本研究中的一些富有攻击性雄鼠身上也能看到。与它们在自然环境中所不同的是，在末端围圈中处于统治地位的雄鼠必须经常与其他雄鼠争斗以保持其地位，且一般常会是几只白鼠打群架。不过，研究者通过观察发现在中间围圈中最强的雄鼠是最正常的。然而，即使是这些白鼠有时也表现出“病理的症状，如发狂；攻击雌鼠、未成年鼠以及不太活跃的雄鼠；并表现出一种特殊的嗜好——一种鼠类的不正常行为——咬其他白鼠的尾巴”（P.146）。

2.服从

与极端攻击行为相反，其他组的雄鼠忽略并避免为争霸主地位而起的战争。其中一组雄鼠是围圈中看起来最健康的白鼠组成，它们身体肥硕、皮毛完好，没有因争斗而常会出现的秃斑。然而，这些白鼠完全不适应社会环境。在围圈之间走动时，它们似乎处于睡眠状态或类似催眠的状态，它们无视其他白鼠的存在，反过来也被其他白鼠所忽视。它们对性活动一点儿也不感兴趣，即使面对发情的雌鼠也无动于衷。

另一组白鼠则极为活跃，并一直在搜寻发情的雌鼠。卡尔霍恩把它

们称为“探索者”（prober）。它们常常被更强势的雄鼠攻击，但他们对争夺地位从不感兴趣。它们的性欲极强，其中相当一部分甚至以同类为食。

3.性偏差

这些探索者也不遵守自然的交配程序。通常情况下，雄鼠追逐发情的雌鼠，直到它逃进洞里，然后，雄鼠耐心地等待，甚至干脆在洞口跳起求婚舞。最后，雌鼠从洞里出来，进行交配。在卡尔霍恩的研究中，大部分进行性活动的雄白鼠遵守这一程序，但探索者例外。它们完全拒绝等待而直接跟着雌白鼠进洞。有时，洞内的窝中还有已死亡的幼鼠，研究者在随后的研究中发现，探索者正是由此开始蚕食同类。

某些特定雄鼠组被称为“泛性者”（the pansexual），因为它们不加分辨地试图与任何其他白鼠进行交配。它们以性为目的，接近其他的雄鼠、幼鼠以及没有发情的雌鼠。这是一个驯服的群体，常常受到更强势的成年雄鼠的攻击，但它们从不为争夺地位而打斗。

4.繁殖异常

白鼠有一种自然本性，那就是筑巢。在本研究中，研究者向白鼠不限量地提供小纸条作为筑巢的原材料。正常情况下，雌鼠非常主动地筑巢，以便为生产做准备。它们将材料集中并堆积起来，形成一个垫子的模样，然后将窝的中间部分做一个凹陷形以放置幼鼠。然而，行为改变的雌鼠丧失了一般的筑巢能力（或倾向），最初它们不能做出中间的凹陷，后来，随着时间的推移，它们收集的纸条越来越少，以至于最终幼鼠直接降生在覆盖着木屑的地板上。

当母鼠预感到存在危险时，它们也丧失了把幼鼠从一个地方转移到另一地方的能力。它们可能转移了一些幼鼠而忘了另一些，或在转移的过程中不小心把它们丢在了地上。通常，这些掉在地上的幼鼠或是被遗弃或是死去，随后被成年鼠吃掉。幼鼠的死亡率在中间两个围圈中最高，约在80%~96%之间。

除此之外，在中部围圈中的雌鼠发情时，它们会被大群的雄鼠追逐，直到它们最终不能逃脱为止。这些雌鼠在怀孕和生产过程中并发症发生率很高，并且它们变得极度不健康。

讨论

你可能会预计这些发现的必然延伸是将它们应用到处于高密度环境的人群中去。然而，由于接下来我们要讨论到的一些原因，卡尔霍恩没

有得出任何这样的结论。事实上，他很少谈论自己的发现——可能的推测或逻辑是：他的结果本身已说明了一切。他只对一个明确的结果进行过评论：白鼠的自然及社会生存行为已被高密度的居住环境所产生的应激明显改变。另外，他提到通过进一步研究、改进方法和修改对结果的解释，他的研究以及其他类似的研究可以有助于我们理解人类在面对类似问题时的情形。

研究发现的意义

就像这本书里的许多研究一样，卡尔霍恩这项研究最重要的一个方面是它引发了许多相关的研究，关注高密度居住环境对人的影响。虽然在这里不可能详细复述所有的研究，但有些例子值得一提。对人来说，促使行为改变发生的环境之一就是特别拥挤的监狱。由国家法官协会资助的一项研究对每个囚犯平均只占50平方英尺[1]的监狱与不太拥挤的监狱进行比较研究。结果发现，前者的死亡率、杀人率、自杀率、生病的次数及惩戒的比率明显较高（McCain，Cox，& Paulus，1980）。但再提醒一下读者，请记住，除了拥挤因素以外，还可能有其他因素影响了这些行为（例如，参见研究37中有关津巴多监狱的内容）。

另一有趣的发现是，拥挤因素对问题解决能力会产生消极影响。有一项研究让人待在非常拥挤的小房间里（每人只有3平方英尺），或待在不太拥挤的较大房间内，并要求被试完成相当复杂的任务，如把不同的形状进行归类，同时给被试听故事，随后要对故事内容进行测试。在拥挤条件下，被试的成绩明显低于不拥挤条件下的被试（Evans，1979）。

请猜一猜，在拥挤环境中你的生理指标将会发生怎样的变化？有研究已显示，你的血压会升高，心率会加快。伴随着这些影响，你会感到其他人对你充满敌意。随着人口密度增加，你会感到时间过得越来越慢（Evans，1979）。

批评

卡尔霍恩对动物的研究结果已得到后继动物研究的支持（参见Marsden，1972）。然而，正如我们在本书的前面所提及的，把动物研究的结果推广到人类时必须特别小心。正如某些物质可能引起鼠类疾病，但对人类的身体健康毫无影响一样；影响鼠类社会行为的环境因素或许不能直接应用于人类，动物充其量只能代表人类的某些方面。有时，动物研究是非常有用的，它可揭示并引导对人进行更加明确的研究；另一些时候，它也可能通向一条死胡同。

1975年，研究者们在纽约城进行了一项研究，试图用人类被试来重复卡尔霍恩的一些研究结果（Freedman，Heshka，& Levy，1975）。对不同居住密度的地区，研究者分别收集了死亡率、生育率（出生率）、攻击行为（法院记录）、心理病理学（精神病院的病人数）等数据。分析所有数据后，研究者没有发现在人口密度和社会病理行为之间有任何显著相关。

虽然如此，在20世纪60年代初期，卡尔霍恩就把工作焦点集中在拥挤情境对人类心理和行为的影响上。因为这与人类相关，这一领域的研究持续至今。

近期应用

约翰·卡尔霍恩在1995年9月7日逝世，他给我们留下了一系列发人深省并具有历史意义的研究。卡尔霍恩在1962年的文章中所论述的各种社会问题与人类生存条件之间表现出愈发紧密的关系。因此，当科学家对攻击行为、不孕症、精神疾病或各种形式的社会冲突等问题进行进一步了解和干预研究时，常会参考卡尔霍恩关于拥挤和行为病理学的研究。

另一项有趣的研究引用了卡尔霍恩的观点，考察在驯养过程中动物行为的改变（Price，1999）。该研究者普莱斯（Price）认为，某些动物被驯化了，即宠物。这些动物经历了若干代遗传与发展的变化，从而改变了自己的行为，这样它们便能与人和平共处，分享同一生活空间。普莱斯的主要观点是，野生动物经过几个世纪的驯养，不得不适应人类的生活方式，虽然这种环境与它们原始的栖息地非常不同。这种生活方式通常包括在相当拥挤的条件下与同类的其他动物、其他种类的动物以及人类和平相处（至少大多数时间是如此）。作者认为，这是通过提高反应阈限的水平而达到的，即要想让一只驯养动物成为占领者和侵略者需要更强的外部刺激。换句话说，在一个相当小的空间中，狗、猫和人类都能够和谐相处，而未经驯养的野生动物间将会逃走和相互厮杀。

一个相关研究发现，与动物相比，人类对人口密度的反应可能存在着一个关键差异。在动物研究中，作为密度增长的直接后果，病理行为是呈线性增长的：一个数据增长，另一个相应跟着增长。然而，雷格兹（Regoeczi，2002）的研究发现，对人类而言，家庭人口密度对增加的社会退缩和攻击性行为的影响实际上是随着一个单独的家庭中人口的增加而降低的。但是，这种影响趋势只有在居住人数没有超过家庭中的房间数时才会出现；当远远超出这个数字时，反社会影响就会伴随着人口密度增加而显现出来。换句话说，当居住条件是5个人居住在一个3居室的公寓里或者7个人挤在一个4居室的屋里时，人们倾向于表现出更多的退缩或攻击性行为。这里有两个可能原因在起作用：也许是人口密度导致了病理行为；也许是那些退缩或攻击性程度高的人出于自愿或被排斥

的缘故，而最终获得不大拥挤的居住环境。

卡尔霍恩的研究有助于精神治疗理论。由于当今世界人口日益增多，人员迁移和多元化发展也更为频繁，对精神治疗领域中提高专业化程度的需求也变得十分必要（Dumont & Torbit，2012）。例如，由于多种文化之间相互作用的增加，对文化交叉相关领域的专业咨询师的需求也逐渐增加。随着寿命的延长，人们会对那些在工作和退休的转型期和慢性病与死亡之间的问题上有专长的咨询师的需求将会增加。随着城市越来越大，越来越拥挤（更多更大的高密度居住区），市中心人口也在不断增多，犯罪率上升。总体来说，很多城市问题会愈发加深，卡尔霍恩的工作与之紧密相关。当这一切发生时，对能够解决各种城市功能障碍的咨询师的需求将会大大增加。

结论

以上这些研究以及其他许多研究表明了社会学家们在探索、完善人口密度和拥挤的影响研究上所作出的不懈努力。尽管社会病理行为产生的原因是繁多且复杂的，然而正是卡尔霍恩45年前的研究第一次让我们注意到了人口密度的影响，这个影响因素虽然只是揭开这个谜团的钥匙之一，但是它所发挥的作用却是至关紧要的。

Dumont, F., & Torbit, G.(2012).The expanding role of the counselor: Fitting means to ends.Canadian Journal of Counseling and Psychotherapy, North America, February 11, 2012.Retrieved from http://cjc.synergiesprairies.ca/cjc/index.php/rcc/artical/view/1776.

Evans, G.W.(1979).Behavioral and psychological consequences of crowding in humans.Journalof Applied Social Psychology, 9, 27–46.

Freedman, J.L., Heshka, S., & Levy, A.(1975).Population density and social pathology: Is there arelationship? Journal of Experimental Social Psychology，11, 539–552.

Marsden, H.M.(1972).Crowding and animal behavior.In J.F.Wohlhill & D.H.Carson (Eds.), Environment and the social sciences.Washington, DC: American Psychological Association.

McCain, G., Cox, V.C., & Paulus, P.B.(1980).The relationship between illness, complaints, anddegree of crowding in a prison environment.Environment and Behavior, 8, 283–290.

Price, E.(1999).Behavioral development in animals undergoing domestication.Applied Animal Behavior Research, 65(3), 245–271.

Regoeczi, W.(2002).The impact of density: The importance of nonlinearity and selection onflight and fight responses.Social Forces, 81, 505–530.

注释

[1]1平方英尺≈0.093平方米。——译者注

第九章 心理治疗

心理治疗（psychotherapy），简单地说，就是“对心理问题进行治疗”。典型的治疗包括治疗者和当事人之间紧密的关爱关系。临床心理学是心理学中关注研究、诊断和治疗心理问题的分支学科。心理治疗的历史主要是由一长串的各种各样的治疗方法组成，每一种方法都被其发明者认为是最有效的。而能证明这些方法有效性的研究并不很充分，也不太科学。然而，这其中产生了一些重要且有影响力的研究突破。

关于心理治疗，常被问及的一个问题是：“哪一种方法最有效？”本章的第一项研究探讨了这一问题，该研究通过一种创新的统计分析方法（就当时而言）证明，各种形式的治疗在一般情况下有效性相等；第二项研究讨论了另一种研究思路，其得出的结果似乎是第一项结论的一个例外；对于恐惧症（对某些事物感到极端的、莫名其妙的紧张）而言，一种称为“系统脱敏法”的行为疗法显然比其他疗法有效得多。该研究由被誉为系统脱敏法创始人的心理学家约瑟夫·沃尔普（Joseph Wolpe）完成；本章的第三、第四项研究涉及两种与心理治疗及诊断相关的诊断工具：罗夏墨迹法和主题统觉测验（TAT），它们常被治疗师用来诊断当事人的心理问题，或是帮助他们探讨其敏感、痛苦甚至是隐秘的心理问题。

研究33 为自己挑选心理治疗师

Smith, M.L., & Glass, G.V.(1977).Meta-analysis of psychotherapy outcome studies.American Psychologist, 32, 752-760.

你不必非得“疯了”才需要心理治疗。大多数人接受心理咨询师和心理治疗师的治疗，不是因为精神疾病，只不过是在生活中遇到了凭借他们通常的应对机制和支持网络不能解决的问题而已。

想象一下，你正处于情绪的低谷，你向周围的密友和家人寻求帮助，但仍不能解决问题。当你忍耐这种痛苦很长一段时间后，你最终决定去寻求专业帮助。因为你是一个消息灵通、敏捷睿智的人，你读过一些心理治疗方面的书，发现有很多可供使用的治疗方法，例如行为疗法（包括研究34中讨论的与沃尔普的研究有关的系统脱敏法）、人本主义疗法、认知疗法、认知行为疗法和种类以弗洛伊德的精神分析理论为基础的心理动力学疗法（psychodynamic therapies）。这些种类的治疗方法虽然使用了不同的基本技术，但其目的都是一样的，即要帮助你改变生活方式，使你成为更快乐、更有创造性和效率更高的人（对各种心理治疗形式更详尽的讨论参见Wood，2007）。

现在，你可能完全糊涂了，如果你需要帮助的话，到底该选择哪种治疗方法呢？下面的内容是你所需了解的：（1）心理治疗真的有效吗？（2）如果它确实有效，那么哪一种治疗方法最好呢？在过去40多年的时间里，心理学家也一直在问同样的问题，知道这一点可能会对你有所帮助（也可能不会）。虽然研究者们进行了许多比较研究，但大多数研究结果都倾向支持进行该研究的心理学家所偏爱的方法。这没什么奇怪的。另外，大多数研究所涉及的被试数量太少，研究手段过于单一。更糟糕的是，这些研究被广泛传播，散见于各种书刊杂志上，因此，要作出完整而全面的判断实属不易。

为了填补心理治疗技术研究文献方面的空白，玛丽·史密斯（Mary Lee Smith）和吉恩·格拉斯（Gene Glass）于1977年在科罗拉多大学开展了一项工作，主要任务是收集截至1977年所有关于心理治疗效果的文献，并重新进行分析。通过对1000多种不同期刊、杂志和书籍的考察研究，他们最后选定了375项研究用来检验心理咨询和心理治疗的效果。之后，他们采用格拉斯发明的元分析统计技术，对所有研究数据进行处理，以从整体上判断不同治疗方法的相对效果（元分析是将许多个别研究的研究结果整合起来进行更大的统计分析，使丰富多样的数据联合在一起，形成一个更有意义的整体）。

理论假设

史密斯和格拉斯的研究目的如下（P.752）：

1.鉴别并收集所有检验心理咨询和心理治疗效果的研究；

2.判断每项研究中治疗方法的疗效如何；

3.比较不同类型心理治疗的效果。

理论假设就隐含在这些目的之中，即通过元分析技术，可显示出心理治疗是否有效，以及不同方法的疗效是否存在明显差异。

方法

虽然史密斯和格拉斯分析的375项研究在使用的研究方法和所评估的疗法类型方面有很大的差别，但每一项研究都至少包括一组接受某种治疗方法的被试与一组接受另一种治疗方法或不接受任何治疗的被试（控制组）。对史密斯和格拉斯来说，对所有研究进行元分析后，各种疗法的有效程度是最重要的发现。效应值从原作者选择使用的治疗方法的所有结果测量中获得。通常，原文会提供一个以上的有效性测量指标，有时候同一指标被使用多次。如自尊心增加、焦虑减少、学业成绩提高和总体适应能力的提高都被用来评估治疗结果。只要有可能，在一项特定研究中使用的所有测量指标，都会包含在元分析中。

从375项研究中共计算出833种不同的效应值。研究涉及约25000名被试。在这些研究中，被试的平均年龄为22岁，他们平均接受17个小时的治疗，治疗师的平均从业年限是3.5年。

结果

第一，史密斯和格拉斯比较了研究中所有接受治疗和没有接受治疗的被试的治疗类型和效果测量指标。他们发现“接受治疗的被试平均恢复水平高于75%的未接受治疗的控制组被试。……有代表性的结果显示，心理治疗将当事人的健康水平从第50个百分位数移到了第75个百分位数”（P.754-755）。百分位数指的是对某一测量指标而言，落在特定分数下的人数在总体中所占的百分比。例如，如果你在一项测验上的得分处在第90个百分位，这意味着与你做同一测验的90%的人分数在你之下。此外，在833种不同疗效中，只有99种（12%）的效应值是负的（意思是说当事人的状态比治疗前更差）。作者指出，如果心理治疗是无效的，那么负效应值应达到或超过50%，即417种。

第二，将各种指标上的所有心理治疗的效果进行比较。结果见图9—1，从图中可清楚地看到，从总体上说，治疗比不治疗明显有效。

第三，史密斯和格拉斯用相似的统计程序比较了所有研究中应用的各种心理治疗方法，图9—2总结了他们关于几种较常见的心理治疗方法的研究结果。

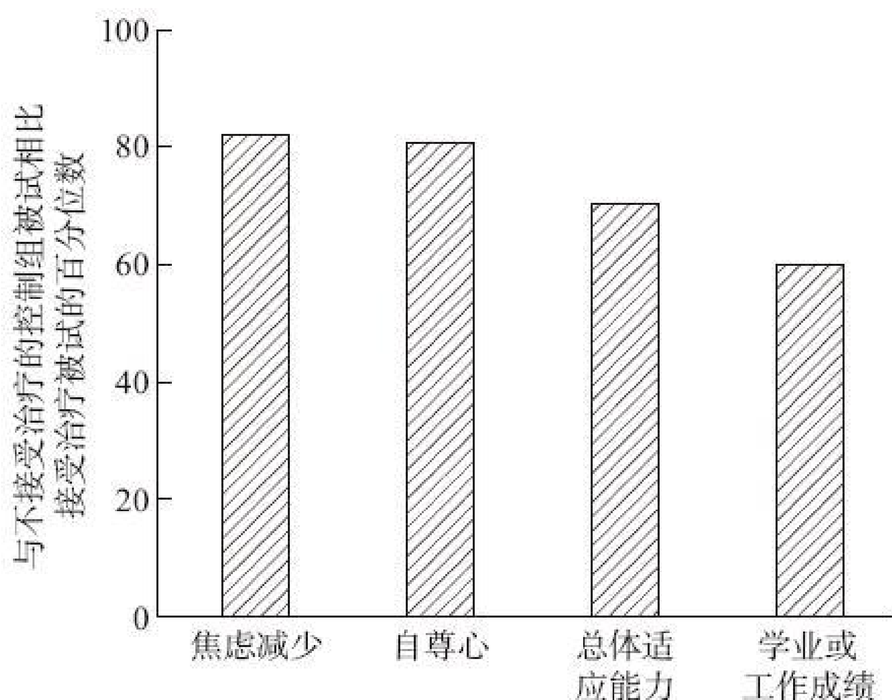


图9—1 从四个方面分析所有研究综合效果的测量

注：如果病人状况没有改善，病人的得分应为50。如果病人的状况变得更糟，病人的得分应低于50。

资料来源：引自P.756的数据。

第四，史密斯和格拉斯把各种治疗方法分成两大类：行为疗法和非行为疗法。其中，行为疗法包括系统脱敏法、行为矫正法和冲击疗法；其余治疗方法均属于非行为疗法。当他们分析用行为疗法和非行为疗法进行的所有研究，并与不接受治疗的控制组比较时发现，两类疗法之间没有差异（相对控制组而言，分别处于第73个百分位数和第75个百分位）。

讨论

总的来说，心理治疗在处理各种心理问题时，都是成功的（见图9—1）。另外，无论不同的治疗方法如何划分归类，它们之间的差异均不显著（见图9—2）。

史密斯和格拉斯从他们的发现中得出三个结论：第一，心理治疗是有效的，元分析的结果清晰地支持了接受心理治疗好于不接受心理治疗的主张；第二，“尽管大量专业人员依据不同的心理学理论在从事心理治疗，但本研究结果显示，不同类型的心理治疗所产生的疗效差异是可以忽略的。认为一种心理治疗绝对比另一种优越……是不恰当的”（P.760）；第三，由于相关信息在各类出版物中的传播还远远不够，研究者和治疗师对于心理治疗效果的假设是不牢靠的。因此，他们提出本研究为解决这些问题迈出了重要的一步，且采用类似技术的研究也值得人们作进一步关注。

恐高症
1. 跨过人行道上的栅栏
2. 坐在三层楼办公室的窗前（不是落地窗）
3. 乘电梯上 45 层楼
4. 看洗窗工人在 10 层楼高的平台上工作
5. 站在椅子上换灯泡
6. 坐在 5 楼房间有栏杆的阳台上
7. 坐在剧院二层前排
8. 站在梯子的第三格上修剪院中的灌木
9. 站在无栏杆的三层建筑的顶部边缘
10. 驾车在盘山道上行驶
11. 像游客一样在盘山道上骑车
12. 站在 20 层建筑的顶部边缘

（引自 Goldstein, Jamison, & Baker, 1980, P.371）

图9—2 七种心理治疗方法效果的比较

注：分数的含义与图9—2相同，分数高于50表明治疗有效。

资料来源：引自P.756的数据。

研究发现的意义和后续研究

史密斯和格拉斯的研究结果减少了当事人对心理治疗效果的困惑，但却增加了心理治疗师的困惑，因为那些选择心理治疗作为职业的人常常会深信某种特殊的治疗方法（他们所选择的）比其他治疗方法更有效。然而，史密斯和格拉斯的研究结果已得到了后续研究的支持（Landman & Dawes，1982；Smith, Glass， & Miller，1980）。这些研究的影响之一是：愿意采用折中疗法来帮助当事人的心理治疗师开始

增多，即在治疗实践中，他们会综合采用几种心理治疗方法，使自己的治疗适合每一位作为特殊个体存在的来访者及其独特的问题。事实上有40%的心理治疗师认为自己采用的是折中法，这个百分比远远高于其他只采用某种特定疗法的医师。

从这些研究以及相关研究中得出“所有的心理治疗方法对所有心理问题和所有的病人都同等有效”的结论，显然是错误的。这些研究从一个非常宏观的角度对心理治疗的疗效进行了综述。然而，依据个人的人格特征及具体问题的具体情况，一些治疗方法可能比另一些更有效。

在你选择心理治疗师的时候，最重要的并不是要考虑他所采用的治疗方法，而是你对心理治疗的期望、心理治疗师的特点以及咨访关系的质量。如果你相信心理治疗会对你有帮助，满怀希望地进入治疗关系，那么治疗的成功概率就会大大增加。你对与治疗师之间关系的感受，也可以使治疗效果有很大的不同。如果你信任你的治疗师，并且相信他或她能真正地帮助到你，那么你的治疗可能会更加有效。

近期应用

史密斯和格拉斯的研究结果和研究方法，在探讨多种形式的心理疗法对各种心理问题的干预效果时，仍有深远的影响。这种影响不仅来自他们得出的结论，即大多数心理疗法的有效性相当，同时也来自于他们对元分析研究技术的应用。

沿用史密斯和格拉斯的方法进行的研究中，有一项研究评估了抑郁症的团体治疗方法的有效性（McDermut，Miller & Brown，2001）。作者对48项以团体形式治疗抑郁症的研究进行了元分析，发现接受治疗的患者通常比未接受治疗对照组中的85%的患者抑郁症状有显著的改善。研究者们推断“团体治疗对抑郁症患者是有疗效的，但是，对于团体治疗可能在哪些方面优于个别治疗则少有实证研究”（P.98）。依据史密斯和格拉斯的研究，你可以预测团体和个别治疗方法的有效性可能是相似的，但需要进一步的研究来明确这一点。

另一项研究展示了史密斯和格拉斯使用的元分析方法的不同应用，该研究关注各种对偏头痛和紧张性头痛患者的行为（即非药物）治疗（Penzien，Rains，& Andrasik，2002）。研究者们用元分析方法比较了30年以来有关放松训练、生物反馈和压力管理三种干预手段的研究。他们发现，总的来说，单独运用行为疗法，偏头痛和紧张性头痛患者的头痛情况就减少了35%~50%。正如作者所指出的，这是一项重要的发现，因为“已有的证据显示，运用行为干预方法改善头痛的水平可以与已被广泛使用的药物治疗相媲美”（P.163）。根据这项发现，作者提出，如果能够使行为治疗更普及、费用更低一些的话，更多的医生和他们的慢性头痛患者可能会选择非药物治疗。

有一项研究证实了史密斯和格拉斯的元分析方法与发现的广泛影响，该研究调查了对精神发育迟滞个体进行心理治疗的有效性（Prout，& Nowak-Drabik，2003）。他们搜集了大量的研究用以进行元分析，这些研究在方法论、心理治疗方式和来访者特点上有着广泛的多样性。分析所有研究得出的结论显示，心理治疗对于精神发育迟滞来访者具有中等程度但十分显著的效应。研究者们推断“应考虑将心理干预手段作为治疗精神发育迟滞者整体计划的一部分”（P.82）。

结论

史密斯和格拉斯的研究在心理学的历史上是一座里程碑，因为这将使人们不再沉迷于试图证明某种具体疗法最为优越，而是把焦点集中在如何更好地帮助那些受心理问题困扰的人们。此后的研究可能更加直接地聚焦于哪些因素能使疗程最短、疗效最令人满意，尤其注重哪些因素能使来访者获得最佳治疗体验。

Landman, J., & Dawes, R.(1982).Psychotherapy outcome: Smith and Glass’s conclusions stand up under scrutiny.American Psychologist, 37, 504-516.

McDermut, W., Miller, I., & Brown, R.(2001).The efficacy of group psychotherapy for depression: A meta-analysis and review of the empirical research.Clinical Psychology: Science and Practice, 8, 98-116.

Penzien, D., Rains, J., & Andrasik, F.(2002).Behavioral management of recurrent headaches: Three decades of experience and empiricism.Applied Psychology and Biofeedback, 27，163-181.

Prout, H., & Nowak-Drabik, K.(2003).Psychotherapy with persons who have mental retardation：An evaluation of the effectiveness.American Journal of Mental Retardation，108, 82-93.

Smith, M., Glass, G., & Miller, T.(1980).The benefits of psychotherapy.Baltimore, MD: Johns Hopkins University Press.

Wood, J.(2007).Getting help: The complete & authoritative guide to self-assessment and treatment of mental health problems.Oakland, CA: New Harbinger Publications.

研究34 缓解你的恐惧心理

Wolpe, J.(1961).The systematic desensitization treatment of neuroses.Journal of Nervous and Mental Diseases，132，180-203.

在我们讨论心理治疗中占据重要地位的系统脱敏疗法（即缓和地逐渐降低你的焦虑或恐惧水平）之前，应首先了解神经症这个概念。神经症这个术语目前已有些过时，它过去常用于描述以极端焦虑为主要特征的一系列心理障碍。现在这些问题统称为“焦虑障碍”（anxiety disorders）。我们都很熟悉焦虑，有时在某些使人紧张不安的场合，如公开演讲、面试、考试，等等，我们会体验到高度焦虑。然而，当某人患有焦虑“障碍”时，他的反应会特别强烈、弥漫、频繁，并导致衰弱。通常这种障碍会影响个体的正常生活，以致其无法发挥正常、应有的生活功能。

最常见的与焦虑相关的障碍有恐惧症、惊恐发作和强迫症。如果你曾经罹患过其中一种疾病，你就会知道这种焦虑可以控制你的生活。这一章中所讨论的约瑟夫·沃尔普（Joseph Wolpe，1915~1997）对这些障碍的治疗工作主要集中在恐惧症上。恐惧症（phobia）一词来源于Phobos这个词，Phobos是希腊恐怖之神的名字。古希腊人把他们想象的Phobos画在面具和盾牌上，用以恐吓他们的敌人。

恐惧症是一种不合情理的恐惧。换句话说，它是一种与实际危险不相称的恐惧反应。例如，当你在森林中的小路上散步，突然一条响尾蛇出现在你面前，盘绕着身子并准备向你进攻，你会感觉到恐惧（除非你是哈利·波特之类的人物）。这不是恐惧症，而是一种正常的、对真实危险合理的恐惧反应。但如果你不敢去动物园，是因为在那儿你可能会看到厚厚玻璃笼中的蛇，那么你就有点儿恐惧症的嫌疑了（除非你是《哈利·波特》中胆小的达力·德思礼）。这在你听起来也许很可笑，但对那些患恐惧症的人来说，一点也不好笑。恐惧症反应是非常难受的，它包括许多症状，如头晕、心悸、乏力、呼吸急促、出汗、发抖、恶心等。恐惧症患者总是警惕地回避那些可能遭遇恐惧刺激的情境，这种回避通常会严重影响一个人正常的生活功能。

恐惧症分为三个主要类型：①单纯性（或特定性）恐惧症，包括对动物不合情理的恐惧（如鼠、狗、蜘蛛或蛇）或对特殊场所，如狭小的空间（幽闭恐怖）或高空（恐高症）的恐惧；②社交恐惧症，以与他人交往时不合理的恐惧为特征，如在公开场合说话，或对困窘的恐惧；③广场恐惧症，是对不熟悉的、开放的、或拥挤的空间的不合理恐惧。虽然各种类型恐惧症的症状是相当不同的，但它们有两个共同的特征：它们都是非理性的，并且都可以用同样的方法进行治疗。

早期对恐惧症的治疗以弗洛伊德精神分析学派的观点为主。这种观点认为恐惧症是无意识心理冲突的结果，这种冲突起源于童年时的精神创伤。这一学派还进一步主张，恐惧症可能是个体不愿面对的某些其他深层恐惧和愤怒的替代品。例如，一个人有不合理的高空恐惧（恐高症），这很可能与他小时候受父亲恐吓有关，他父亲假装要把他从悬崖上推下去。作为一个成人，要承认这种经历可能迫使他去接受曾遭父亲虐待的事实（他不想面对的一些事），于是他压抑了这些事实，并以恐惧症的形式取而代之。根据对问题根源的弗洛伊德式的看法，精神分析一直试图通过帮助病人了解无意识感情，释放隐藏的情感，在此过程中使他们摆脱恐怖。虽然有时这种技术在治疗其他类型的心理障碍时很有帮助，但在治疗恐惧症时却效果不佳。研究发现，有时即使某人揭示了无意识中与恐惧相关的心理冲突，恐惧症仍然存在。

虽然，约瑟夫·沃尔普不是第一个提出使用系统脱敏法这一行为技术的人，但他被公认为是对此技术进行改良，并用来治疗焦虑障碍的第一人。行为疗法与精神分析最大的不同是，它不关心心理障碍的无意识根源和受到压抑的冲突。行为治疗的基本观点是：你习得了一种无效的行为（恐惧症），现在你必须消除它。这便是沃尔普治疗恐惧症的基础。

理论假设

沃尔普和其他人在早期研究中已发现，动物的恐惧反应可以通过一个简单的条件反射程序而得到减轻。例如，一只老鼠看到一幅逼真的猫的照片时，会表现得非常恐惧。但如果在每次呈现猫的照片时，都给老鼠食物，老鼠的恐惧会越来越小，直到最后恐惧反应完全消失。老鼠最初把猫的图片与恐惧条件反射性地联系在一起。然而，老鼠进食所引起的反应与恐惧反应是不相容的。由于恐惧反应和进食反应不能同时并存，所以前者被后者抑制了。这两种不相容的反应称为交互抑制（当两种反应彼此抑制，在既定时刻，只有一种反应可以存在）。沃尔普提出了更普遍的假设，即“如果一个反应在引起焦虑的刺激呈现时能抑制焦虑的产生……存在于刺激和焦虑之间的联结将被削弱”（P.180）。他还认为，人类的焦虑反应与那些在动物实验室里发现的焦虑反应非常相似，因此交互抑制的观点也可用来治疗人类的各种心理障碍。

在对人的治疗中，焦虑的抑制反应是深度放松训练而不是进食。这一观点的理论基础是：你不可能在同一时间里既体验到深度的身体放松，又体验到恐惧。作为一名行为主义者，沃尔普相信患恐惧症的原因是你通过经典条件反射过程，在生活的某个时间点习得了恐惧。通过经典条件反射，某些物体与极度的恐惧在你脑中形成联系（参见研究9中巴甫洛夫的研究）。从华生（参见研究10中关于小艾尔伯特的内容）及其他学者的研究中，我们知道这种学习甚至可以在年幼时发生。为了治疗你的恐惧症，当你面对一个恐怖情境时，你必须经历一种能抑制恐惧或焦虑的反应（放松）。这种治疗技术有效吗？沃尔普的研究报告是基

于从150例恐惧症患者中随机抽取的39例被试，他用系统脱敏技术对这些被试进行了治疗。

方法

假设你患有恐高症，且这一问题非常严重，以至于你很难爬上梯子去修剪你庭院里的树木或走上二楼的办公室，你的恐惧症严重干扰了你的生活，于是你决定寻找像沃尔普这样的行为治疗师进行心理治疗。你的治疗将包括放松训练、焦虑层次的建立和脱敏三部分。

放松训练

在最初的几次治疗中，对恐惧症的针对性治疗非常少。相反，治疗师把精力集中在教你如何放松你的身体上。沃尔普采用了一种由埃德蒙·雅克布森（Edmund Jacobson）在1938年发明的渐进性肌肉放松练习，这种放松练习在当前心理治疗中仍经常使用。该过程包括拉紧和放松你全身的各组肌肉（如双臂、双手、面部、后背、腹部、双腿等）以达到深度放松，放松训练可能要在治疗师指导下进行数次，直到你能独自达到放松状态。通过这种训练，无论何时，你都可以让自己处于放松状态。为了确保完全的放松，沃尔普对许多被试配合使用了催眠术，但现在已明确这点对有效的治疗通常是不必要的，因为人们往往不需要催眠术也可达到完全的放松。

焦虑层次的建立

治疗过程的下一阶段是要你和你的治疗师共同制定一张焦虑产生情境或恐惧产生场景的条目表。条目表以引起轻微不适的情境开始，强度逐渐增加，直到你能想象到的能引发最严重焦虑的事件。层次的数量从5或6，到20或更多。表9—1列出了可能的恐高症层次表，还有一张直接来源于沃尔普研究报告的幽闭恐惧症患者的层次表。

表9—1 焦虑层次

恐高症
1. 跨过人行道上的栅栏
2. 坐在三层楼办公室的窗前（不是落地窗）
3. 乘电梯上 45 层楼
4. 看洗窗工人在 10 层楼高的平台上工作
5. 站在椅子上换灯泡
6. 坐在 5 楼房间有栏杆的阳台上
7. 坐在剧院二层前排
8. 站在梯子的第三格上修剪院中的灌木
9. 站在无栏杆的三层建筑的顶部边缘
10. 驾车在盘山道上行驶
11. 像游客一样在盘山道上骑车
12. 站在 20 层建筑的顶部边缘

（引自 Goldstein, Jamison, & Baker, 1980, P.371）

幽闭恐惧症
1. 阅读矿工被困的文章
2. 不用清洁剂清除指甲油
3. 被告知某人在监狱里
4. 拜访某地后不能离开
5. 手指上有一枚戴得很紧的戒指
6. 在乘坐火车的路途中（路途越长、焦虑越严重）
7. 与电梯工作人员一起乘电梯（乘电梯时间越长，焦虑越严重）
8. 独自乘电梯
9. 所乘的火车穿越隧道（隧道越长，焦虑越严重）
10. 被锁在房间里（房间越小，时间越长，焦虑越严重）
11. 被困在电梯里（时间越长，焦虑越严重）

（引自 Wolpe, P. 197）

脱敏

现在，进入实际的“消退”阶段。根据沃尔普的观点，不直接与恐惧情境接触，对减少病人对情境的敏感性是必要的（这是来访者很乐于听到的）。相同的效果可以通过描述和想象来完成。记住，恐惧症是你通过联想而产生的。因此，你将以同样的方式消除恐惧症。首先，治疗师将要求你进入深度放松状态，就像你曾学过的那样。然后，治疗师便从焦虑层次的第一条开始，给你一一描述情境：“你走在人行道上，来到一个大栅栏前。当你继续往前走，你可以看到栅栏有4英尺高。”你的任务就是想象这个情境，并同时保持完全放松。如果成功，治疗师将进行下一个步骤：“你坐在三层楼的办公室内……”等等。如果在此过程中的任何时候，你感到轻微的焦虑，就举起你的食指示意，这时治疗师将暂时停止对焦虑层次表的呈现，直到你再次完全放松为止。然后当你能保持放松状态的时候，描述重新从前面某一个条目开始。这个过程一直持续到你完成全部层次表中的条目并能保持放松状态。一旦你完成了这个过

程，在后面的治疗期内你可以多重复几次。在沃尔普的研究中，病人治疗成功所需要的次数因人而异。有些病人不超过6个疗程就康复了，而有一个病人治疗了近100个疗程（这是一个对死亡有严重恐惧的病人，而且还患有另外两种恐惧症）。平均疗程在12个左右。这个疗程数比正规精神分析一般需要的时间要少得多，精神分析治疗通常需要持续一年。

与此治疗方法相关的重要的问题是，系统脱敏疗法真的有效吗？

结果

在沃尔普的文章中报告的这39个病例患有多种不同类型的恐惧症，包括：与他人在一起、幽闭恐怖、暴风雨、被注意、拥挤、亮光、伤口、恐高症、摔倒、拒绝和蛇形物。判断治疗成功与否，是根据病人的自我报告和现场的直接观察。通常，病人以描述脱敏过程的方式来报告病情好转和逐渐康复，沃尔普以此描述为依据来判断病人。他把脱敏过程评定为完全成功（完全脱离恐怖反应）、部分成功（恐怖反应只有原来的20%或更少）和不成功三个等级。

在这39例患者中，共有68种恐惧症状需要治疗。在这些治疗中，其中62种症状（涉及35例患者）的治疗被评定为完全成功或部分成功，因而成功率为91%，其余6种症状（占9%）的治疗没有成功。成功治疗的平均疗程为12.3次。沃尔普的解释是：大多数不成功的病例存在特殊的问题，例如，不能想象层次表中呈现的情境，因而无法进行适当的脱敏。

对沃尔普的批评主要来自于弗洛伊德的精神分析学派阵营，他们认为，沃尔普的方法只针对了症状而并没有针对引起焦虑的根源。他们认为以这种方法治疗的症状消失后，一定会出现新的症状来取代它们。他们把这比喻为一个有漏洞的大堤，一个洞被堵住了，另一个漏洞最终又出现了。沃尔普通过不同时期的跟踪报告，对这些批评和质疑作出了回应，他对接受治疗并获得成功的35例病例中的25例进行了4年多的跟踪。通过对这些报告的研究，他写道："没有一例病例报告复发，也没有出现新的恐惧症症状或其他神经症症状。当系统脱敏治疗完成后，我再也没有观察到神经焦虑的再次出现，或许事实就是如此吧！"（P.200）

讨论

在沃尔普的文章中，他把讨论的焦点集中在反击精神分析学家对他的怀疑上。在20世纪50年代，精神分析仍是一种非常普遍和流行的心理治疗形式。当行为疗法刚刚开始进入临床心理学的主流时，它引发了大

量争论。沃尔普指出，脱敏疗法与传统的心理分析相比有以下几大优势（参见研究报告原文，P.202）：

1.在每一个案例中，心理治疗的目标都可被明确地陈述出来；

2.焦虑的来源可被清楚且迅速地确定；

3.在根据焦虑层次表对场景进行描述时，病人反应上的变化在治疗的每一时期都可测量。

4.治疗可以在有其他人在场的情况下进行（沃尔普发现，在治疗期间如果有他人，如见习治疗师在场，并不会影响疗效）；

5.如果愿意或需要，可更换治疗师。

后续研究和近期应用

自从沃尔普发表了这篇文章并出版了一本关于在心理治疗中应用交互抑制方法的书（Wolpe，1958）以来，系统脱敏疗法的应用发展到了顶峰，现在它常被认为是治疗焦虑障碍，特别是恐惧症的首选。这种发展大部分应归功于对这一方法疗效更新、更科学的研究。

保罗（Paul，1969）完成了一项著名研究，他针对那些对公开演讲感到特别恐惧的大学生进行了治疗。首先，要求所有被试在一名不熟悉的听众面前做一小段即兴演讲。以观察者的评定、生理指标和自陈问卷来测评被试的焦虑程度。然后，将被试随机分配到三个不同的治疗组：（a）系统脱敏治疗组；（b）顿悟治疗组（与精神分折相似）；（c）无治疗组（控制组）。有经验的治疗师将所有治疗分为5个疗程，然后将所有被试安排在相同的公开演讲场合，并测量所有焦虑指标。图9—3展示了研究结果。显然，在所有治疗方法中，系统脱敏治疗对减少焦虑明显更有效。更具有说服力的是，两年的跟踪调查显示，系统脱敏组中85%的人仍有明显改善，而顿悟组只有50%。

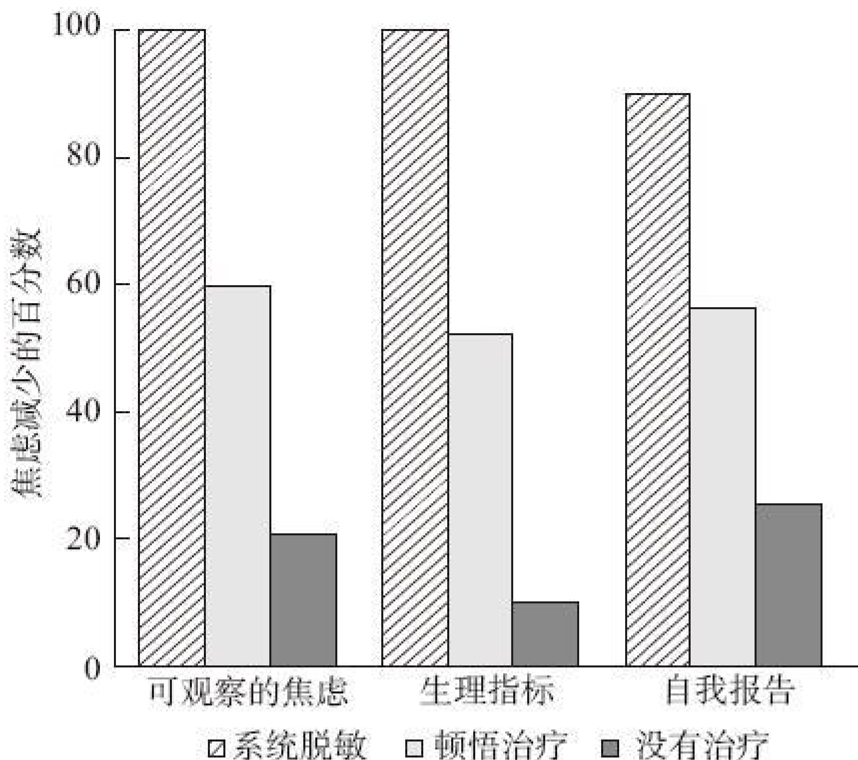


图9—3 焦虑治疗的结果

资料来源：引自Paul，1969。

众多有关行为治疗的研究仍然引用沃尔普早期的著作作为他们理论支持的一部分。他将经典条件反射概念应用在心理疾病治疗中，使之成为普遍情况下干预策略的一部分。例如，有一项研究（Fredrickson，2000）以沃尔普的交互抑制概念为理论依据，发展了一种新的治疗策略，专门治疗由焦虑、抑郁、攻击和应激等消极情绪引起的健康问题。弗里德克森（Fredrickson）提出，应帮助和教会那些有这类心理问题的病人产生更多和更强烈的积极情绪，如爱、乐观、快乐、兴趣和满意，这将直接抑制消极思维。他认为：

积极情绪通过消除由特殊因素引起的不良身心状态，来解除消极情绪对个体思维和机体的束缚……治疗使人的健康状态更趋完善，并培养了人的积极情绪。这种积极情绪不仅可抵制消极情绪，而且拓宽了个体惯常的思维模式并建立了个人的应对策略资源（P.1）。

另一篇以沃尔普的研究为依据的文章，探讨了系统脱敏疗法治疗数学焦虑症的有效性（Zettle，2003）。对于这种恐惧症，相信很多学生都不会陌生。在这项研究中，沃尔普的治疗技术被用来帮助学生克服极端的数学焦虑水平。研究者给被试提供渐进式肌肉放松训练的指导语和一盘每天在家练习放松的磁带。研究者和每一名学生共同商讨列出一个

有11项条目的数学焦虑层次表，包括诸如“被数学老师点名在黑板上解题”或者“在期末考试中遇到一道不会做的应用题”这样的条目（P.205）。然后，如前面所描述的，给每一名学生呈现焦虑层次。简言之，的确有效！治疗结束时，12名学生中有11名“显示他们的数学焦虑水平恢复正常或得到改善……而且，在治疗结束后的两个月间，对被试数学焦虑的治疗仍显示出显著的效果”（P.209）。

结论

沃尔普在他的文章中随即指出，克服恐惧和焦虑的想法并不新鲜。“其实人们很早就知道，增加与恐惧物体的接触机会会使恐惧渐渐消失”（P.200）。事实上，即使你在读这一项研究之前，从来没有听说过系统脱敏治疗，但这些道理也可能是你早已知道了的。例如，假设一个13岁的男孩有一种对狗的恐惧，这种恐惧可能是由于在男孩更小的时候，曾有过与狗有关的可怕经历，如被大狗攻击、被任何一只狗咬伤，甚至是父母亲中有一人怕狗（通过模仿机制，恐惧症可以由父母传递给孩子）。由于这样的经历，男孩在狗与恐惧之间形成了联系。如果你想治疗男孩对狗的恐惧，你会怎么切断这种联系呢？多数人对这个问题的第一反应是：“给这孩子买一条小狗！”如果你也是这么想的，那你就使用了系统脱敏疗法。

Fredrickson, B.(2000).Cultivating positive emotions to optimize health and well-being.Preventionand Treatment, 3 (article00001a): 1–25.Retrieved February 3, 2008, at http://www.unc.edu/peplab/publications/cultivating.pdf

Paul, G.L.(1969).Outcome of systematic desensitization: Controlled investigation of individualtechnique variations and current status.In C.Franks (Ed.), Behavior Therapy: Appraisal andStatus.New York: McGraw-Hill.

Wolpe, J.(1958).Psychotherapy through reciprocal inhibition.Palo Alto, CA: Stanford University Press.

Zettle, R.(2003).Acceptance and commitment therapy (ACT) vs.systematic desensitization intreatment of mathematics anxiety.Psychological Record, 53，197–215.

研究35 投射出真正的你

Rorschach, H.(1942).Psychodiagnostics: A diagnostic test based on perception.New York: Grune & Stratton.

在一个温暖的夏日，你和你的朋友躺在一片绿油油的草地上休息，蓝天上飘着几朵棉絮般的白云，你指着其中一朵白云对朋友说："看，那朵白云多像一个穿着婚纱的新娘，拖着长长的裙摆。"你的朋友回答说："哪儿？我看不出来。在我看来，那倒像一座火山，有一缕轻烟从它的顶部升起。"当你们试图劝说对方接受彼此对于同一形状的不同知觉时，气流的变化又将云朵变成了完全不同的样子。但是，为什么你们俩看到的东西竟然会如此不同？你们面对的是同样的形状，而对它的解释却是两种完全不相关的物体。

由于每个人的知觉都受心理因素的影响，因此从云的形状里看出来的不同物体或许代表观察者人格的某些方面。换句话说，就是你和你的朋友把自己心里的某些东西投射到了天空中的这些形状上。这就是赫尔曼·罗夏（Hermann Rorschach，1884~1922）构建其"图形解释测验"（form interpretation test）的理论基础，也叫"墨迹测验"（the inkblot test），它是投射技术这种心理测验工具的最早版本之一。

最著名且应用最广泛的两个投射测验是罗夏墨迹测验（本研究中将讨论）和主题统觉测验（简称TAT，见研究36）。这两个测验在临床心理学的历史上占有重要位置。罗夏墨迹测验在1921年第一次出版，其中涉及对多组精神病患者的直接比较，常用于有关心理疾病的诊断。

投射测验就是给被试呈现一个模棱两可的图形，并假设被试在描述这个图形时会把自己内部的或无意识的心理过程投射其中。在罗夏墨迹测验中，刺激只不过是一幅对称的墨迹图，它是如此模棱两可，以至于几乎可以被知觉为任何事物。罗夏认为，一个人把墨迹看成是什么东西常常在很大程度上揭示了他或她真实的内部心理加工过程。他称之为"对随机图形的解释"（the interpretation of accidental forms）。有一个常被提到的有关罗夏墨迹测验的故事，讲的是一个心理治疗师给一个病人做罗夏墨迹测验，在呈现第一张图片时，心理治疗师问："这卡片使你想起了什么？"病人回答："性。"呈现第二张图片后，治疗师问了同样的问题，病人又回答："性。"当第一组的5张图片都得到了同一个答案时，心理治疗师评论道："看来，你的大脑已经被性占据了。"对此，病人吃惊地回答道："我？医生，是你呈现了这些有关性的图片！"当然，这个故事把罗夏墨迹测验过分简单化了，而且通常来讲，性方面的含义不应该比其他方面更多。

罗夏相信，他的投射技术可以用作两个主要目的：一是它可以作为一种研究工具，以揭示人格的无意识方面；二是罗夏稍后提出的，他希望这个测验可用于各种类型的精神病理学诊断。

理论假设

罗夏技术的理论基础是，在对随机墨迹进行解释时，被试的注意力被转移，使得其惯常的心理防御相对减弱。进而，深藏于心中的想法慢慢浮现。倘若被知觉的刺激是模棱两可的（也就是说，没有多少线索能说明它是什么），那么对刺激的解释就必将来自于个人的内心深处（与此概念相关的讨论参见研究36中默瑞的TAT测验）。在罗夏的概念中，墨迹要尽可能地模棱两可，才能使人最大限度地投射自己的无意识。

方法

罗夏形成墨迹测验的方法可分为两个主要部分：一是他编制原始图形的过程。二是对被试或当事人的反应进行计分和解释的方法。

测验的发展

罗夏对于墨迹图形如何制作的解释，听起来很像是创作有趣的儿童艺术作品："制作这种随机图形非常简单：把少许大滴墨水滴在一张纸上，将纸对折，墨水就会在纸的两半对称地扩散。"（P.15）然而，简单的部分到此为止。罗夏进一步解释说，只有那些满足一定条件的图形才能使用。例如，图形必须相对简单，并对某些含糊的物体有适度的暗示。他还提出，图形必须是对称的，因为不对称的墨迹会由于难以解释而为被试拒绝。经过大量测试，罗夏最终确定了10个图形，组成了最早的罗夏墨迹测验。其中，五张图是黑白色，两张是黑色和红色，三张是彩色。图9—4呈现了罗夏所用图片类型的四张图形。

施测和计分

罗夏图形解释测验的施测非常简单，即每次呈现给被试一张图片，并提问："这可能是什么？"被试可自由地改变图片的方向，也可以按自己的喜好将它放得离眼睛远点儿或近点儿。施测人员或治疗师不给被试任何提示或建议，而是只记录被试对每个图形的所有反应。测验没有时间限制。

罗夏指出，被试几乎都会认为这个测验是用来测试人的想象力的。然而，他非常认真地解释说，它不是一个想象力测验，并且人想象力的

创造性高低不会明显改变测验的结果。罗夏认为，这是一个知觉测验，并涉及感觉、记忆、无意识和意识的过程，这些刺激图形与个体内心的某种力量联系在一起。

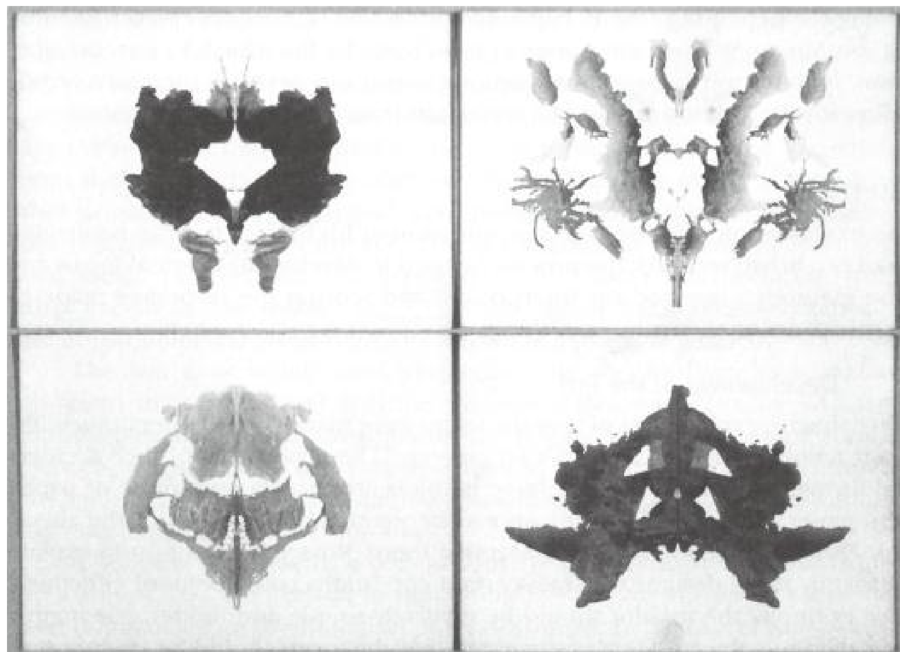


图9—4 与罗夏图形解释测验中使用的图片类似的一些随机图形示例

资料来源：Science Museum/SSPL/The Image Works。

罗夏列出了被试对10张墨迹图回答的几条计分规则（P.19）：

1.总共作出多少次反应？反应时间，即在回答之前被试花了多长时间看图片？被试拒绝解释卡片的次数有多少？

2.被试的解释只取决于图形的形状？还是也包含了颜色？或在知觉中是否包含了运动？

3.图形被看成是一个整体还是几个部分？哪些部分被分离出来？被试对此如何解释？

4.被试看到了什么？

有趣的是，罗夏认为在被试对墨迹做出的回答中，对图形内容的解释是最不重要的。以上我们总结了罗夏依据这四条计分规则，对众多有不同心理症状的被试进行观察的结果。

结果

为了解不同人群在墨迹测验中可能做出的不同表现，罗夏和他的同事对来自不同群体的被试进行了施测。这些群体不仅包括受过不同教育的正常个体，还包括精神分裂症患者及躁狂抑郁症患者。

表9—2列出罗夏所报告的对10张墨迹图的典型反应类型。当然，这些反应会因个体或群体的不同而发生变化。在表中列出的仅是一些例子。

表9—2 一般正常被试对墨迹图的典型反应

图形数	反应
Ⅰ	两个圣诞老人用手臂夹着两把扫帚
Ⅱ	一只蝴蝶
Ⅲ	两个提线木偶的形象
Ⅳ	家具上的一件装饰品
Ⅴ	一只蝙蝠
Ⅵ	一只飞蛾或一棵树
Ⅶ	两个人头或两个动物的头
Ⅷ	两只小熊
Ⅸ	两个小丑或燃烧的火焰
Ⅹ	一只兔子的头，两条毛毛虫或两只蜘蛛

（引自P.126-127上数据）

罗夏发现被试通常对10个图形共给出15到30个反应。抑郁的被试一般给出的答案较少；快乐的被试给出的答案较多；精神分裂症病人的答案数目在个体间有很大的差异。全部测验通常需要20分钟~30分钟完成，精神分裂症患者所用时间平均较正常人少很多。正常被试几乎从不拒绝对所有图形做出回答，但精神分裂症患者常常拒绝回答。

罗夏认为，被试关注图形的哪一部分、解释中是否包含运动、反应中有多大程度涉及颜色等对分析被试的测验结果都是非常重要的。这些常常比被试具体看到了什么更重要。他关于这些因素的计分方案是非常复杂的。而且临床医生需要经过训练和体验，才能熟练恰当地分析被试的回答。不过，格莱特曼（1991）已经为我们提供了关于计分过程的有效而简洁的总结：

> 使用整个墨迹图进行解释显示了被试综合思维和抽象思维的能力，而解释细节部分的频率很高则意味着被试有强迫性的固执。相对较多地使用白色部分是叛逆和消极的信号。描述人在运动的回答显示了被试的想象力和丰富的内心生活；在回答中颜色占主导地位说明被试较感情用事和冲动（P.684）。

罗夏发现，关于被试实际上在墨迹中看到了什么，最常见的回答为动物类和昆虫类。回答中涉及动物的反应比例占25%~50%。有趣的是，抑郁症患者回答动物的比率最高，而艺术家回答动物的比率却最

低。

罗夏提出的另一类型的反应是“独创性回答”（original responses）。这指的是那些在每100次测验中发生的概率少于1次的回答。独创性回答在被诊断为精神分裂症的被试中最为常见，而在中等智力水平的正常被试中最为少见。

讨论

在对图形解释测验所进行的讨论中，罗夏指出，测验的最初目的是为研究一些有关人脑和心理无意识方面的理论问题，而发现该测验有作为诊断工具的潜力纯属偶然。罗夏认为，他的测验常常可以预示精神分裂症倾向、隐匿性神经症、潜在的抑郁、内向或外向性格和智力。但他的意思并不是说墨迹测验能取代通常的临床诊断技术，只是可以对诊断过程有所帮助。罗夏也警告说，虽然测验可显示一定的无意识倾向，但不能详细探明无意识的内容。他认为其他当时的常用心理学技术，像弗洛伊德梦的解析和自由联想是解释深层无意识的更好方法。

批评和后续研究

自从罗夏编制了墨迹测验后，几十年以来大量研究都对罗夏的许多结论提出了质疑。最重要的批评之一与测验的效度有关——测验是否测量到了罗夏所要测量的东西，即潜在的无意识的心理问题。有研究证明，许多罗夏归因于心理人格反应差异，可以很容易地为下列因素所解释：如言语能力、被试的年龄、智力水平、受教育程度甚至是施测者的人格特征（关于这些问题更详细的讨论参见Anastasi & Urbina，2007）。

科学研究显示，从整体上讲，罗夏墨迹测验作为一种人格测验或诊断工具，其准确性并不令人满意。尽管如此，该测验在临床心理学家和心理治疗师中仍应用甚广。这种明显的矛盾可以这样解释：罗夏墨迹测验技术确实常在临床中应用，但它并不作为正式测验使用，而只是作为治疗过程中，治疗师增加对当事人的了解以及打开话题开始交流的一种手段。其实质是延伸了治疗师和当事人之间的言语交流。测验中对反应的用途较少受严格限制，一些治疗师认为这为有效的心理治疗提供了有帮助的见解。

近期应用

对最近的心理学相关文献进行回顾发现，罗夏评定量表的效度仍然是人们研究和争论的焦点（关于这一争论的全面概述参见Woodet al.，2003；Exner & Erdberg，2005）。从精神分析研究前沿得到的几项研究结果表明，新的施测和计分方法可提高量表的评分者信度，且可增进诊断和鉴别各种心理障碍的能力。例如，阿瑞拉和（Arenella）和奥杜夫（Ornduff，2000）采用罗夏墨迹法研究受性虐待女孩和其他应激情境中未受虐女孩在在身体意象方面的不同。研究者发现，前者在罗夏墨迹测验中的作答方式表明她们比后者更关注自己的身体。类似地，研究者得到了一组66名年龄在14~17岁之间的年轻的精神病男性罪犯的罗夏墨迹测验分数（Loving & Russell，2000）。该研究发现，至少罗夏墨迹测验的某些标准变量与精神病理学的各种水平存在明显联系。作者认为，罗夏墨迹测验可以预测哪些青少年正处于暴力犯罪的高危阶段，因此有助于人们进一步改进预防和干预策略。

一项研究推进了关于罗夏墨迹测验效度争论的有趣进展，该研究在对性犯罪者的心理评估中，将罗夏墨迹测验和客观性心理测验的明尼苏达多相人格调查表MMPI（Minnesota Multiphasic Personality Inventory）相比较（Grossman等，2002）。在评估性犯罪者是否患有心理障碍时存在一个常见的问题，那就是他们经常会否认有心理问题或将任何心理问题的严重性降到最低。这项研究发现，能够在MMPI测验中“装好”而得出正常的心理剖面图的性犯罪者，在罗夏墨迹测验中会作为心理异常者被曝光。“这些发现显示罗夏墨迹测验能够察觉‘装好’的企图，并因此可能在常有罪犯故意歪曲事实的法庭上提供有价值的信息。”（P.484）当然，经过改编的罗夏墨迹测验的效度与原始的罗夏墨迹测验存在同样的问题。

结论

这些研究连同许多其他的研究均证明了罗夏所作贡献的持久影响力。伴随着罗夏墨迹测验的改进和更广泛的应用，未来的研究可能会指向投射测验的开发与改进，使其能够提供更好的科学效度和更有价值的治疗洞察力。

Anastasi, A., & Urbina, S.(2007).Psychological testing, 7th ed.Upper Saddle River, NJ: Prentice Hall.

Arenella, J., & Ornduff, S.(2000).Manifestations of bodily concern in sexually abused girls.Bulletin of the Menninger Clinic, 64(4), 530–542.

Exner, J., & Erdberg, P.(2005) The Rorschach, advanced interpretation.Hoboken, NJ: Wiley.

Gleitman, H.(1991).Psychology, 3rd ed.New York: Norton.

Grossman, L., Wasyliw, O., Benn, A., & Gyoerkoe, K.(2002).Can sex offenders who minimize onthe MMPI conceal psychopathology on the Rorschach? Journal of Personality Assessment, 78, 484–501.

Loving, J., & Russell, W.(2000).Selected Rorschach variables of psychopathic juvenile offenders.Journal of Personality Assessment, 75(1) , 126–142.

Wood, J., Nezworski, M., Lilienfeld, S., & Garb, H.(2003).What's wrong with the Rorschach? Scienceconfronts the controversial inkblot test.New York: Wiley.

研究36 编个故事吧

Murray, H.A.(1938).Explorations in personality (PP.531-545).New York: Oxford University Press.

在研究35中，我们结合罗夏墨迹测验介绍了一种叫作“投射测验”的方法，它常被一些临床心理学家用来揭示被试的某些人格特征。罗夏墨迹测验背后的思想就是让个体把自己的解释投射到客观上毫无意义、也没有任何结构的图形上。为了获知被试的人格特征，罗夏考察了被试对墨迹图中的特定部分及其各种具体细节的关注和对图形中运动的感知，从而得出关于被试人格特征的结论。虽然该测验也考虑被试解释的情节内容，但这不是最重要的。

在罗夏墨迹测验问世几年以后，亨利·A·默瑞（Henry A.Murray，1893~1988）和他的同事克里斯蒂安娜·摩尔根（Christiana D Morgan）在哈佛心理诊所合作编制了另一种完全不同的投射测验，称为“主题统觉测验”（Thematic Apperception Test，TAT）。该测验把注意的焦点全部集中在被试解释的内容方面。与罗夏墨迹测验无结构的图形不同，TAT测验由一套黑白图片组成，图片中人物所处的情境没有特定含义。该测验要求心理治疗中的当事人或被试看图来讲故事，然后治疗师或研究者将故事的内容进行分析，以期揭示被试隐藏在无意识中的冲突。

TAT测验的理论基础是，当你观察人的行为时，无论是在图画中，还是在现实生活中，你对该行为的解释将以情景中可获得的线索为依据。倘若行为的原因显而易见，那么你的解释将不仅仅是大致正确的，而且也会与大部分观察者相一致。然而，如果情景模棱两可且很难找到行为的原因，那么你的解释似乎将更多地反映出与你自身有关的某些东西——你的恐惧、愿望、冲突等。例如，假设你看到一个男人和一个女人的脸，他们仰望着天空，表情迥异：男的似乎受到惊吓，而女的正在开怀大笑。你觉得难以解释他们的表情，然而一旦进一步仔细观察，你发现他们正排队等着乘坐“京达卡”（Kinda Ka），它位于美国新泽西州六旗大冒险乐园，是世界上最高、最快的云霄飞车。现在，你觉得在该情景中推测这对夫妇的行为更为容易了，并且你的分析可能与其他观察者更为相似。现在假设你在孤立的、没有任何线索的情景下看到相同的表情，然后问你：“这些人在干什么？”你的回答将以你的内心想法为解释，并且它会更多地揭示你的内心世界而不是你观察的对象。此外，由于孤立行为的意义模棱两可，因此不同观察者的回答将有很大区别（例如，他们正在看飞碟，或在看一场滑雪赛，或在看孩子们在玩攀高游戏，抑或是看一场逼近的暴风雨）。这些个人认知看法的差异形成了默瑞和摩尔根的主题统觉测验的理论基础。

理论假设

TAT测验的基本假设与罗夏测验相同，即认为人的行为是由无意识力量驱动的。这一观点隐含着最早由弗洛伊德发展起来的心理动力学的原理（参见研究30中有关弗洛伊德理论的论述）。该理论认为，为了准确诊断和成功治疗心理问题，必须揭示无意识冲突（通常形成于童年）。这是研究35中讨论过的罗夏墨迹测验的目的，它同样也是默瑞TAT测验的目标。

默瑞写道："这一程序的目的就是刺激文学创作，从而激发幻想，以揭示在潜意识中的情结。"（P.530）他对这个过程的构想是，给被试呈现模棱两可的描绘人类行为的图画。为了解释图画，当事人的自我意识会有所下降，也不太会意识到治疗师正在观察他们。这将进一步减少当事人的自我防御，揭示出其内心的愿望、恐惧以及被压抑的过去经历。默瑞还指出，这一测验的理论基础的一部分是"大量的文学创作是作者自身的经历或幻想在意识或潜意识中的表现"（P.531）。

方法

TAT测验的最初设想是，让被试猜想是什么原因导致了画中描述的情景，其结果将会如何（如图9—5所示）。测验的结果显示，比起让他们猜想这幅画所描述的事实，让被试简单地就这幅画编一个故事能使主试获得更多的信息。



图9—5 一张TAT图片的样例。你会怎样解释这张图画呢？

资料来源：Van D.Bucher /Science Source。

莫瑞和摩尔根研发了一些图画，来激发人们产生关于其生活中的内心冲突和重要经历的幻想。因此，他们决定，每一幅图画必须至少包括一个易为被试辨认的人物。通过对几百幅图画的反复试误，最终选择了20幅图画。由于TAT测验现在已为人们熟知，许多人认为测验用图片的广泛传播会降低测验的效度。不过，如果你没有见过测验所选择的图画类型，就很难理解该测验。因此，图9—5呈现了最初备选图中的一幅图画，但它没能入选最终的20幅测验图画中。

默瑞对TAT测验进行了一项早期研究，并将研究结果收录在自己1938年出版的一本书中。这项研究的被试均为年龄在20～30岁之间的男性。每位被试坐在一张舒适的椅子上，背对着主试（这是心理治疗师施测TAT测验时的常见情景），以下是给每个被试的确切指导语：

这是一个测量你创造性想象力的测试。我将呈现给你一幅图画，希望你根据这幅画，编一个情节或一个故事。画中人物的关系如何？在他们身上发生了些什么事？他们当时的想法和感受是怎样的？结果会如何？请尽你的最大努力来完成这项任务。我要求你尽情发挥你的文学想象力，因此，你尽可以按照你希望的长度和详细程度编写这个故事（P.532）。

主试依次向被试呈现图画，并记录被试对每幅图画的描述，每个被试有一个小时的时间。由于时间的限制，大多数被试只能完成20幅图画中的15幅。

几天以后，被试被重新召回，并被问及与故事有关的一些问题。为了掩饰研究的真实目的，被试被告知该研究的目的是比较他们与著名作家的创造力。主试重复被试上一次对图画的反应，并要求他们解释故事的来源。而后让他们接受一个自由联想测试，在这个测试中主试要求被试在主试报出一串词以后，说出自己的第一反应。这些工作是用来判断被试依据图画所编的故事在多大程度上反映了他们的个人经历、冲突、愿望，等等。

结果和讨论

默瑞和摩尔根报告了他们从这项早期研究中所得出的两个主要发现。

第一是发现被试所编的故事有四个来源：（a）书籍和电影；（b）发生在朋友和亲属身上的真实生活事件；（c）被试自己的亲身经历；（d）被试意识或潜意识中的幻想（P.533）。

第二项发现，也是更重要的发现，即被试很清晰地把自己的人格、情感以及心理投射到他们的故事之中。作者报告的其中一个例子是，大多数学生被试会把一幅画中的主人公看作学生，而非学生被试就不会这么做。另一个例子中，一名被试的父亲是一名船工，并且被试有一种想去周游世界的强烈愿望，这种幻想出现在他对好几幅图画的解释中。

例如，对一幅描绘两名工人谈话的图画，这位被试的故事是："这两个小伙子是一对冒险家，他们总是筹划在不同寻常的地方会面，他们现在正在印度，并已听说南美发生了一场革命，他们在商讨如何到那里去……最后，他们搭乘货船到达了那里。"（P.534）默瑞在报告中写到，无一例外，每一个参加研究的人都把自己人格的某方面投射到他们的故事中。

为了进一步说明TAT测验是如何反映人格特征的，作者对其中一个被试的情况做了详细报告。维尔特（Virt）在第一次世界大战期间，经受了迫害、饥饿、与母亲分离等这些恐怖的童年经历后，从俄国移民到了美国。默瑞这样描述TAT测验的第13张图画（这里没有画出）："一个

男孩儿坐在地板上，背靠着长沙发，头斜靠在右臂上，卷曲着身子。在他旁边的地板上有一个形似左轮手枪的东西”。（P.536）当维尔特看到这幅画时，他所编的故事如下：

可怕的事情发生了，他所爱的某个人——也许是他的妈妈自杀了。她可能因为贫穷而自杀。他已经是一个懂事的孩子，明白眼前发生的一切，他也想自杀。但他毕竟是一个孩子，并且没多久就振作起来。有一段时间，他一直处在不幸之中，最初的几个月他想到死（P.536）。

把这个故事与最近其他人就同一幅画所编的故事进行比较，你会有非常有趣的发现：

1.一名35岁的初中老师描述道：“我想这幅画的主人公是一个没有犯罪却被判坐牢的人。他否认自己犯过罪并在法庭上一遍又一遍地抗争。但最终他放弃了。现在，他已经精疲力竭、抑郁而绝望。他做了一支玩具手枪企图逃跑，但他知道这也没有用。”（引自作者的文件）

2.一名16岁的中学生描述道：“这个女孩可能正与她的兄弟玩捉迷藏游戏，她正在从1数到100。她很忧伤也很疲惫，因为她从来都没赢过但又不得不玩。在此之前男孩子们似乎在这儿玩过其他游戏，因为有一把玩具枪在那里。”（引自作者的文件）

你不必像一名心理治疗师那样去预测这三个人可能把他们的哪些内部冲突、动机或愿望投射到了这一幅图画上。这些例子也证明，对于TAT测验，不同的人可能会有明显不同的回答。

批评和相关研究

虽然TAT测验所使用的刺激与罗夏墨迹测验差异很大，但它也在信度和效度方面遭到同样的批评（参见研究35中罗夏墨迹测验中对该问题的有关论述）。TAT测验最严重的信度问题是，对同类反应，不同的临床医生会给予不同的解释。有人认为，治疗师可以在不经意间把自己的潜意识特性加入到被试对图画的描述之中。换言之，对TAT的解释可能会变成对实施该测验的临床医生的投射测验！

就效度（即TAT对其所要测量的东西测到什么程度）而言，人们会提出数种批评。如果该测验测量了基本心理过程，那么它就应该可以鉴别正常人和精神病人，或鉴别不同类型心理疾病。但研究显示，它不能进行这样的鉴别。在爱伦（Eron，1950）的一项研究中，对两组男性退伍军人实施TAT测验，一组是大学里的学生，而另一组是精神病院的病人。在对TAT测验结果进行分析后，研究者发现在这两组之间或在不同类型的精神病人之间均无显著性差异。

其他一些研究对TAT测验预测行为的能力提出了质疑。例如，如果一个人在其故事中涵盖了大量的暴力，这并不能区分这种攻击性只存在于其幻想中，还是真实暴力行为的潜在表现。有些人容易产生攻击性的幻想而从不表现出实际的暴力行为，而另一些人的攻击性幻想则预示着实际的暴力冲动。由于TAT测验的反应不能说明某一个体应归为哪一类，因此该测验在预测攻击倾向方面的价值很小（见Anastasi & Urbinai，1996）。

另一种对TAT测验基本的、也是非常重要的批评（批评内容之一也指向罗夏墨迹技术）是对投射假说本身的有效性提出怀疑。TAT测验依据的假设是，人们对图画所作的描述揭示了他们的基本人格特征，以及持续进行的潜意识心理过程。不过，有科学证据表明，对像罗夏墨迹测验和TAT测验等投射测验的反应，会更多受随机因素和情境因素的影响。这就意味着，假设你周一刚做完工作、与老板大吵一架后做一次TAT测验，然后周六从海滨轻松地玩了一天回来再做一次TAT测验，在这两种情况下，你依据图画所编的故事可能会完全不同。批评者指出，造成故事内容不同的原因是因为TAT测验只反映了你当前的心理状态，而不是你的真实的潜在自我。

作为该批评的一项证据，有研究已找到了各种与TAT测验结果相关的影响因素：饥饿、缺乏睡眠、药物使用、焦虑水平、挫折、语言能力、施测人员的特征、被试对测验情境的态度和被试的认知能力。依据这些发现，心理测验领域的权威人物之一安妮·安娜斯塔西（Anne Anastasi）曾写道："许多研究倾向于质疑投射假说，有充足证据表明两选一的方式能更好地解释个体对无结构刺激的反应。"（Anastasi & Urbinai，1996）

近期应用

每一年，默瑞的研究和主题统觉测验仍然被众多的人格特征研究及人格测验的研究所引证和使用。一项研究将诊断为分离性障碍的患者，如创伤性遗忘和分离性身份识别障碍（以前被称为多重人格障碍）与精神病院的住院病人的TAT作答进行比较（Pica et al.，2001）。研究者们发现，在分离性障碍病人中，TAT图片的作答几乎不包含积极情绪。"分离性障碍被试的测验行为特征包括转换、恍惚状态、访谈中的遗忘（在访谈过程中对部分内容的遗忘）和对图片产生强烈的（激动的）抵制情绪"（P.847）。

默瑞1938年的研究也为不少有关人格障碍的研究所借鉴，这些研究涉及反社会人格（不顾及别人的权利；缺乏内疚感或良心）、回避型人格（长期的和持续的不恰当情绪）、边缘型人格（激愤，极不稳定的人际关系）和自恋型人格（极端的自我重要感，非常需要别人的赞美）。一些研究发现，TAT和测验可成功鉴别所有这些人格障碍并且TAT分数

与MMPI（明尼苏达多相人格问卷）分数一致。MMPI是一个应用广泛并有较高效度的较客观的人格评定工具（Ackerman et al.，1999）。

人们对含义模糊图片的解释绝大多数是随着文化的不同而改变的，承认这一点非常重要。有研究证实了这一点。该研究分析了赞比亚青少年与德国青少年对TAT的反应（Renaud & Murry，1996）。如你所想，这两个被试组在文化、信仰、受教育程度和生活经历方面都非常不同。作者发现，两个被试组对研究中所使用的五张TAT图片的意象复杂性和解释有非常显著的差异，以至于作者指出，事实上使用TAT这个方法比较不同文化下人们的重要心理变量可能是无效的。

结论

像TAT测验和罗夏墨迹测验等投射测验最引人注目的一个方面是，尽管有大量的证据谴责它们信度、效度低，理论基础可能不正确，但它们仍在治疗师常用心理测验之列。事实上，临床医生仍然热衷于这些测验工具，而实验心理学家却渐渐对它们有了戒心，这成为两组不同专业的人群争论的焦点（关于此问题的综述，参见 Lilienfeld，Wood，& Garb，2000）。我们如何调和这一矛盾呢？对这一问题最普遍的答案是，TAT和罗夏墨迹测验通常用于心理治疗而非用作正式的诊断工具，更多的是作为临床医生与其病人早期访谈的补充。随后，治疗师们以非常个人化的方式使用这些测验以打开其与来访者的交流通道，并进入来访者若没有TAT的故事提示可能回避或隐藏的心理领域（参见Cramer，2006）。正如一名执业心理治疗师所说：“虽然我不给当事人对TAT的回答打分或用它做诊断依据，但是图画是一种奇妙而有效的工具，把我们带入当事人生活中受困扰的领域。这些问题从进行TAT测验的过程中流露出来，能使治疗更有针对性，也更有效。”（引自作者的文件）

Ackerman, S., Clemence, A., Weatherill, R., & Hilsenroth, M.(1999).Use of the TAT in the assessment of DSM-IV Cluster B personality disorders.Journal of Personality Assessment, 73(3), 422-442.

Anastasi, A., & Urbinai, S.(1996).Psychological testing, 7th ed.New York: Macmillan.

Cramer, P.(2006).Storytelling, narrative, and the Thematic Apperception Test.New York: GuilfordPress.

Eron, L.(1950).A normative study of the thematic apperception test.Psychological Monographs, 64 (9, Whole No.315).

Hofer, J., & Chasiotis, A.(2004).Methodological considerations of

applying a TAT-type picture-story test in cross-cultural research: A comparison of German and Zambian adolescents.Journal of Cross-Cultural Psychology, 35(2), 224–241.

Lilienfeld, S., Wood, J., & Carb, H.(2000).The scientific status of projective techniques.Psychological Science in the Public Interest，1, 27-66.

Pica, M., Beere, D., Lovinger, S., & Dush, D.(2001).The responses of dissociative patients on the TAT.Journal of Clinical Psychology, 57, 847-864.

第十章 社会心理学

社会心理学是心理学的一个分支，它主要探讨的问题是：别人怎样影响你的行为，而你又怎样影响他人的行为。简言之，社会心理学就是研究人与人相互作用的一门科学。心理学的这一分支内容很庞杂，范围涉及广泛，包括浪漫的爱情关系、集体行为，以及偏见、歧视、攻击行为等多种主题。这可能是很多非心理学家的人认为与自己个人生活最为相关的一个心理学研究领域。我们所有人大多数的清醒时间都在以不同方式与他人进行相互作用，所以我们自然想知道更多包含在我们的社会关系之中的心理学过程。而且，社会心理学包含大量具有里程碑意义的科学研究，其数量之多在心理学领域中也屈指可数。

本章所选取的四项研究以以下几种方式使心理学发生了明显的变化：（1）为人们提供了看待人类社会行为的全新视角；（2）激起新的研究浪潮，使人们或证实，或改变，或反对原有的研究结果；（3）引起了关于研究伦理的热烈争论，并最终促成了在本书前言中讨论的伦理守则。

第一项研究是心理学史上最为著名的研究之一：菲利普·津巴多的“斯坦福监狱实验”，这项研究产生了一些关于监禁心理学的惊人启示；第二项研究探讨了从众效应对行为的决定作用；第三项研究揭示了一个令人惊讶的现象，叫作“旁观者效应”，这种现象是指在危难时刻，旁观者越多，愿意伸出援手的人就越少；最后，我们所要接触的也许是心理学史上最闻名于世的研究（在某种意义上也可以说是名誉尽失的研究）——斯坦利·米尔格拉姆（Stanley Milgram）的“盲目服从权威的研究”。

研究37 斯坦福监狱实验

Zimbardo, P.G.(1972).The pathology of imprisonment.Society, 9(6), 4–8.

Haney, C., Banks, W.C., & Zimbardo, P.G.(1973).Interpersonal dynamics in a simulated prison.International Journal of Criminology & Penology，1, 69–97.

你曾经被监禁过吗？先假设你我的回答都是“没有”。那你认识在监狱里待过的人吗？可能吧。无论如何，我们大多数人都对在监狱中生活所产生的心理效应知之甚少。你可能读过一些关于监狱的文章、故事或者小说，而且几乎可以肯定你曾经在电影或电视中看到过描述监狱生活的场景。基于这些了解，大多数人都可以确定监狱一定不是我们愿意卷入的地方！我们知道坐牢是一种可怕的经历，犯人会产生一些激烈反应，甚至是病态的行为。而且我们中的大多数人还认为那些被选作监狱工作人员的人，例如看守和狱警，可能拥有一些独特的人格特征。但是，行为科学家们如何系统地研究对于监狱工作人员和犯人来说，监狱经历所产生的心理和情绪效应呢？

正如大多数复杂的现实生活情境，研究监狱生活的心理学对于研究者来说是个挑战，因为使用的方法必须是相关研究。这就是说，我们可以观察监狱环境，访谈犯人和看守，在犯人释放后收集他们的信息，然后试图基于这些分析形成假设。但是我们不能科学地控制监狱环境，因而对于我们所观察到的行为的真正原因无法得到清晰有效的结论。是监狱改变了人们，还是在监狱中的人原本就是与常人“不一样”？一种绕过这个研究困境的方法可能是创建一个模拟的“科研监狱”，然后把人们分为“囚犯”和“看守”置于其中。听起来不可能吗？如果现在来做一项这样的研究可能会很困难，但是在30年前，著名的心理学家菲利普·津巴多和他的同事克雷格·哈尼（Craig Haney）、柯蒂斯·班克斯（Curtis Banks）、戴夫·杰夫（Dave Jaffe）在斯坦福大学进行了这样的研究（在本节开头列出的两篇文章就是关于他们研究的最早论述）。他们想创建一个模拟的监狱，其中由一些大学生随机分配扮演“看守”和“囚犯”。他们的“监狱”建在斯坦福大学校园里心理学系大楼的地下室（有关细节稍后会详述）。

理论假设

津巴多在研究中验证的假设是，相对于你的内在性情和天性，周围的环境和情境会更强有力地影响你的行为。津巴多主张，尽管我们拥有某些特定的遗传或内在行为倾向，但是强大的环境会战胜这些内在倾

向，并导致我们做出一些与平时的自己截然不同的行为。津巴多和他的同事们想要看看，如果把正常人置于一个对个人施加强大影响的情境中，即监狱中，会发生什么情况。

除了环境对我们的行为具有强大影响这一最初的信念以外，研究者们并没有建立其他特定的假设。为了检验情境力量的影响，他们将所有被试随机指派为“看守”或者“囚犯”。这种随机的指派会导致两种角色在模拟监狱环境中表现出显著不同的反应，例如，在交往方面的行为测量、在心境和病理学方面的情绪测量、对于自我的态度以及应对和适应这一新异情境的一些其他指标（Haney，Banks, & Zimbardo，1973）。

方法

设置

津巴多的目标是创设一个尽可能类似于监狱的情境。他请来一位曾坐过15年牢的前罪犯作为顾问。尽管在这项研究中监狱并不是真实的，而且研究中的被试都知道这一点，但是津巴多希望确保这个过程模拟了真实的监狱体验。



图10—1 “斯坦福监狱”的一个典型“牢房”

资料来源：菲利普·津巴多公司。

津巴多监督着一组工作人员将斯坦福大学心理学系大楼地下室的一些房间和走廊改造为一所“监狱”。由于研究计划持续两个星期，所以这所监狱需要好好建设。走廊的两头都拓宽了，实验室的房间成为监狱的牢房。为了增强真实感，单个牢房的门都在小窗上装了竖栏并配有单独的牢房号码（见图10—1）。沿着牢房的封闭走廊是“监狱院子”，囚犯被试可以从他们的牢房里出来到那里吃饭和活动。在走廊的尽头，是一间设计成单独囚禁的小黑屋，用来禁闭那些惹是生非、反抗和不尊重看守的，或者其他一些不合作的囚犯。浴室就在大厅里，但是守卫会把囚犯的眼睛蒙上把他们带过去，所以囚犯们并不知道他们的方位（Zimbardo，2007b）。“监狱”里装了一个隐藏的摄像机和内部通话系统，这样实验人员就可以保持对看守和囚犯行为的监督。

被试

如果你之前不熟悉这个研究，那么你接下来读到的内容可能会让你觉得很惊讶甚至是震惊。在你往下读的时候，试着把自己置于那些被试的心态。首先，研究者在位于斯坦福大学附近的加利福尼亚州帕洛阿尔托县的当地报纸上刊登广告，以每天15美元的报酬来招募志愿者参与有关监狱生活的实验研究（现在相应的被试报酬大概是每天75美元）。为了保证被试知情同意，研究者告诉这些志愿者该研究的大体性质，而且告诉他们在研究中可能会有一些违背他们个人隐私和公民权利的情况，他们获得的食物可能仅仅达到满足基本营养需要的最低限度。所有的志愿者都同意了以上各项条款。

研究者进行了大量的测试以排除有心理问题和犯罪前科的志愿者，最终从将近100名志愿者中选取了24名大学生。之后，通过投掷硬币的方式随机将这些人分成了“看守”和“囚犯”两组。要记得，津巴多在这个研究中的目标是将决定行为的内在人格因素和外部情境影响分离开来。所以，有必要确保这两组被试在开始的时候是尽可能同质的（Zimbardo，2005）。然后，没有任何的指导语、训练、或是对于接下来研究的准备，所有的被试都直接回到了自己的家。

研究程序

本研究的目标是观察、记录和分析囚犯和看守的行为。正如之前提到的，津巴多和他的同事们在寻找一些迹象，来说明这些年轻人所置身的情境和角色强度足以胜过他们每个人的人格特征和行为倾向。

“囚犯”：在初步筛选几天之后的一个星期天早上，分派到囚犯组的被试在自己家中被“突然袭击”，帕洛阿尔托警察局的一名（真正的）警官出人意料地敲响了被试的家门。所有的被试都以“抢劫罪”被逮捕，所有人都被搜查过、戴上手铐，被响着警笛、闪烁着警灯的警车迅速带离。所有囚犯都被登记在册、采集指纹，然后被蒙着眼睛关到单人牢

房。之后，他们被告知将要被转移到“斯坦福县监狱”（也就是心理学系大楼地下室的那个模拟监狱）。

当这些囚犯到达监狱后，那些被分派充当守卫的被试对他们进行了搜身（见图10—2），脱光他们的衣服并用气雾喷雾器进行除虱。守卫们给每个“囚犯”发了以下物品：一件类似于罩衫的囚服，囚服上分别有四位数字的号码（这些号码将成为囚犯们在整个研究期间的代号）；橡胶拖鞋；需要一直套在头发上的尼龙袜（用来模拟大多数监狱中给囚犯剃光头的样子）；缠在脚踝上的加锁铁链（铁链并没有真的锁在什么东西上，只是为了提醒囚犯的身份）。津巴多指出，尽管这些程序与现实中的真实程序相比有所变化，但是这些程序背后的思想就是模拟羞辱、镇压，以及在真实监狱中囚犯的体验。这些囚犯每三个人被分在一个小囚室中，每人有一张简易床和一个薄床垫、一条毯子。三张简易床填满了整个空间，囚室里再没有空余的地方了。

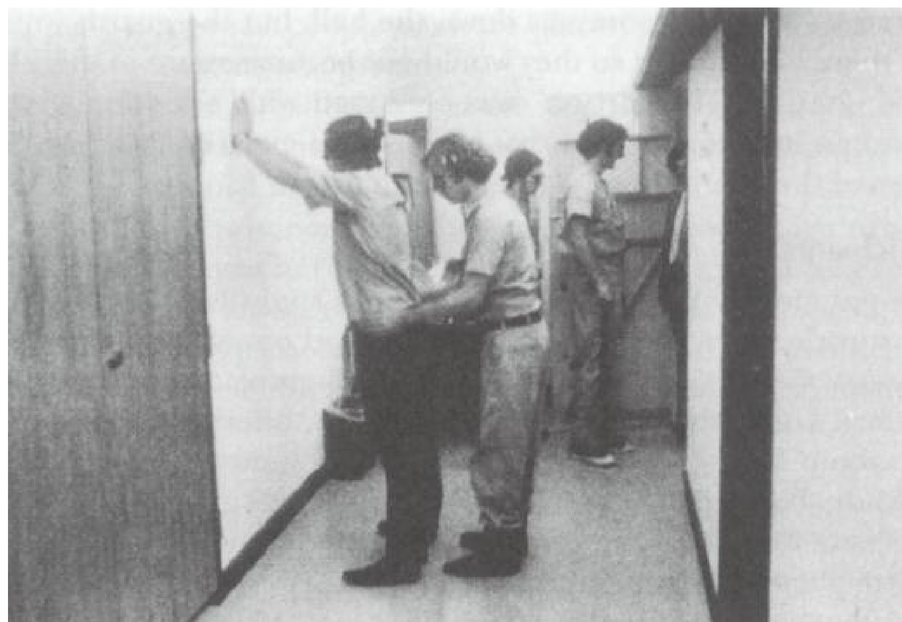


图10—2 斯坦福监狱的“守卫”在对新来的“囚犯”搜身

资料来源：菲利普·津巴多公司。

“守卫”：与囚犯不同，守卫们不需要一直待在监狱里（毕竟他们没有被监禁），他们分为每组三个人，每组值8个小时班，不值班的时候他们仍然过自己的生活。他们领到了相同的监狱守卫制服、警棍（虽然不容许他们击打犯人）、反光墨镜（让他们看起来更有威慑力，而且外貌不太明显）。津巴多解释说他关于反光墨镜的想法来自保罗·纽曼（Paul Newman）在1967年主演的一部电影《铁窗喋血》（Cool Hamd Luke）（Zimbardo，2007）。这些守卫没有受过任何关于他们所扮演的角色的专门训练，而且仅仅被告知他们的责任是让囚犯们守规矩和维

持监狱的秩序。

结果

这是心理学史上被最多探究、讨论和分析的研究之一。守卫们和囚犯们身上发生的人格和行为变化是如此地意义深刻且令人震惊。为了在有限的篇幅内总结那些复杂的研究发现，在表10—1中简要列出了被试的一些典型行为。在这里更笼统地描述了接下来的几天里“斯坦福监狱”中所发生的事情。

出乎任何人的预料，这些囚犯和守卫身上的真正特性和人格似乎很快就消失了，取而代之的是他们被要求扮演的角色。在一天之内，“演戏”和现实生活之间界限就变得模糊，令人感到不安。正如津巴多在他的最初的研究中关于这些被试的描述（1972）：

> 大多数人真的变成了“囚犯”和“守卫”，不再能区分角色扮演和自我……在不到一周的时间里，监禁的体验（暂时地）破坏了一生的学习；人类的价值被搁置，自我概念受到挑战，人类本性中最底层、最丑陋的病态面显露出来。我们觉得非常恐怖，因为我们看到一些男孩们（守卫）把其他人当作是卑劣的动物一样对待并且乐于享受那些残忍的行为；而另一些男孩们（囚犯）变成了卑屈顺从、失去人性的机器人，他们只能想到逃跑、自己个人的生存，以及对守卫们越来越多的仇恨（P.4）。

要记得，这是一个由高水平的专业研究者开展的一项科学研究，而且它很快就有了自己的生命力。参与的被试们，尤其是那些扮演囚犯的被试，似乎忘记了他们是具有自由意志的大学生。他们随时可以直接退出研究，但是并没有这么做。在几天以后，许多被试申请假释或释放，但是当释放请求被驳回后，他们仅仅是沮丧而顺从地回到牢房。囚犯组的被试中有五个人的情绪崩溃和应激反应非常强烈，他们变得抑郁、无法清晰思考甚至绝食。因此，研究者不得不在研究开始的最初几天就让他们退出了研究（或更确切地说，是让他们离开了这个监狱）。

其中一些守卫开始折磨囚犯，显然他们很享受由自己的地位所带来的权力。有些守卫不那么严厉，并且努力去保持公平，但是他们也一直没有干涉那些残暴专横的守卫，而且更重要的是，他们也从未提示实验者其他守卫可能超过了其角色的分寸。甚至津巴多本人也时常忘记他在负责一项科学研究，而是发现自己不自觉地进入“监狱主管”的角色。

表10—1 “囚犯”们和“守卫”们在“斯坦福监狱”研究中的行为和反应

守卫	囚犯
对囚犯使用一些贬损的、侮辱性的语言；骚扰和恐吓囚犯	很快变得温顺、卑屈，遵守守卫们提出的规则
对囚犯进行羞辱性的评论（例如，"2354 号囚犯，你过去告诉 2578 号囚犯你爱他。"）	表现出明显的创伤和抑郁的早期迹象，包括哭泣和深深的沮丧
（每天）半夜粗声喊醒所有囚犯以清点人数	乞求假释
经常把做俯卧撑作为对轻微过错的惩罚（一个守卫在囚犯罚做俯卧撑的时候踩在他的背上。）	愿意放弃实验报酬来换取释放
似乎很享受对囚犯们的虐待性控制	
为了镇压叛乱，用灭火器向囚犯们喷射干冰	情不自禁地哭泣、愤怒、思维混乱
整夜地关押囚犯单独囚禁	筹划并实施了一次"叛乱"，脱掉丝袜头套、撕去囚犯上的编号、用床把牢房挡起来、诅咒和辱骂守卫
把去卫生间作为一种特权，不时地拒绝囚犯上厕所的申请而在牢房里放置一只垃圾桶	
在牢房里安置线人（实验者的助手），用于监视犯人们逃跑迹象或叛乱迹象	设计了详尽的逃跑计划，但一直没有实施
在囚犯们的逃跑计划暴露后脱光他们的衣服；移除囚犯们的床并强迫他们交出毯子	最终完全放弃了反抗和相互支持
随机地给某些囚犯特权（更好的食物、可以刷牙、洗澡，等等），以求破坏囚犯们之间的友情、信任和团结并对其各个击破	树立了"人人为己"的态度，并放弃与其他囚犯之间的团结
为了惩罚囚犯们的逃跑尝试，强迫他们用手清洗马桶、将"夜间点名"延长几个小时、增加罚做俯卧撑的数量	随着研究的进展，日益绝望地温顺地接受了守卫们的贬低和虐待
不断想出新的方法来打击囚犯们的意志	在 6 天之后，所有的囚犯都变得消极被动、失去人性、如同机器人一般

（Haney et al.，1973; Zimbardo，1972；Zimbardo，2005；Zimbardo，2007b.）

近期应用

如同米尔格莱姆关于服从的研究（详见研究40）一样，津巴多开展的监狱研究在这三十多年来产生了全面的社会影响和政治影响。即便这不是不可能，但是要脱离研究的政治属性来讨论津巴多的研究结果是一件困难的事情。在美国以及世界上其他大多数国家，监狱改革都是最受争议的热点话题之一。在历史上，针对囚犯的系统化虐待是有文件记录的，而且这一情况延续至今。从津巴多开展研究的那个年代到今天，美国关于监狱暴动、叛乱、反叛、绑架及谋杀的历史充满了大量与斯坦福大学地下室所发生的那些事件类似的事情。美国监狱和拘留所中囚犯的数量从20世纪80年代的50万左右增加到2006年的220万人，这进一步加剧了虐囚事件发生的潜在可能（Bureau of Justice Statistics，2007）。

美国在监人员的数量位于世界之首。而且，20世纪70年代中期以来，监狱的教育改造目标基本上被废弃了（尽管罪犯教养所的名称还在广泛使用），取而代之的是惩罚目标和将罪犯与公众隔离的目标（称为剥夺其能力）。在1998年，津巴多和哈尼分析了自从他们在斯坦福的研究以来，监狱系统是如何变化的。这里呈现的是津巴多关于他们分析结论的表述：

监狱采用惩罚和将罪犯隔离的内在模式，而不是其他一些可能会降低高累犯率的基本改造措施，这仍然是失败的社会实验。我们的分析显示，监狱的条件在我们研究之后的几十年里显著恶化，这是监狱政治化的后果。媒体的夸大其词让选民们担心犯罪，而政治家、检察官、地方检察官及其他一些官员则对犯罪采取强硬立场，以此来获取选民的支持（Zimbardo，2005）。

阅读到这里，你可能在思考津巴多的监狱研究与过去15年中美国参与的中东战争中发生的一些事件之间的可能联系。一些高度报道的事件，尤其是伊拉克阿布·格莱布监狱虐囚丑闻和关于古巴关塔那摩湾拘留营虐待拘留者的报道（参见Hooks & Mosher，2005；Keller，200），将“斯坦福监狱研究”重新带回公众关注的焦点。津巴多在其最近的著作《路西法效应：理解好人是如何变坏的》（The Lucifer Effect.Understanding How Good People Turn Evil，2007a）中，回顾了监狱研究并把他关于囚犯虐待的研究和评论从监狱拓展到人类罪恶这一更大的概念。我们不相信类似阿布格莱布监狱事件可能会真实地发生，不相信任何人，尤其是自由、民主社会中的公民可能会对其他人做出如此残酷的虐待。怎么会发生这样的事情呢？像津巴多一样的心理学家和其他社会科学家们试图帮助我们理解这种现象。正如一位研究这些虐待事件的作者指出：

新闻记者们在寻找一些社会科学研究来理解在伊拉克、阿富汗以及世界上的其他地方发生的虐待事件。这些报告不再强调少数的“烂苹果”，转而质疑对于那些最底层士兵的惩罚。津巴多（1972）的研究在这些报告中很引人注意。他不同意“烂苹果”理论，而是主张是装苹果的木桶坏了。津巴多批评布什政府的“领导失败”，他着重指出那些对囚犯的虐待性审讯手段和严厉待遇得到了布什政府最高级官员和军队官员“由上而下的授权”（Hooks & Mosher，2005，P.1632~1633）。

从来自伊拉克、阿富汗、关塔那摩的一份份报告当中，我们听闻并从图片细节中看到守卫和审讯官对囚犯实施的可怕的虐待和折磨。而在所有的报告中，这些守卫和审讯官们就像津巴多监狱实验中的被试们一样，并不是虐待成性的、残忍的人。他们在本质上就是普通人，可能和你我并没有什么不同，而他们却发生了戏剧性的转变。这可能最终源于所有罪恶中最强大的情境力量——战争。

结论

正如之前提到的，津巴多原计划进行两周的研究，但是他在第六天的时候就终止了实验，因为这个模拟的监狱情境太强大了，而且以惊人

的方式变成了现实。实验人员和随机指派的大学生们不复存在，他们逐渐成为自己所扮演的角色，变成囚犯、守卫和监狱长。这些角色的力量是那么地强大，以至于个人身份消失了，被试甚至是实验人员都难以意识到在“斯坦福”监狱中的行为是多么危险。以下是津巴多对于他叫停研究的决定的表述：

我终止实验不仅仅是因为暴力的持续升级和“守卫们”与“囚犯们”的敌对恶化，……而且是由于我自己意识到我自己所经历到的个人的转变。……我自己变成了监狱的负责人，而研究负责人的功能被置于第二位。我开始像一位严厉的机构负责人一样说话、走路和行动，我更多地考虑“我的监狱”的安全，而忽视了作为一位心理学研究者应当照顾的年轻人的需要。在某种意义上，我认为关于这个情境的力量的最深刻的评估正是它改变我自己的程度（Zimbardo，2005，P.40；及Zimbardo，Maslach，& Haney，1999）。

Bureau of Justice Statistics (2007).Number of persons under correctional supervision.Retrieved February 4, 2008, from http://www.ojp.usdoj.gov/bjs/glance/tables/corr2tab.htm

Haney, C., Banks, W.C., & Zimbardo, P.G.(1973).Interpersonal dynamics in asimulated prison.International Journal of Criminology & Penology，1, 69–97.

Hooks, G., and Mosher, C.(2005).Outrages against personal dignity: Rationalizing abuse and torture in the war on terror.Social Forces, 83(4)，1627–1645.

Keller, A.S.(2006).Torture in Abu Ghraib (Iraq prisoner abuse scandal, 2004).Perspectives in Biology and Medicine, 49(4), 553–569.

Zimbardo, P.(2005).A situationist perspective on the psychology of evil: Understanding how good people are transformed into perpetrators.In A.Miller (Ed.), The social psychology of good and evil: Understanding our capacity for kindness and cruelty (pp.21–50).New York: Guilford.

Zimbardo, P.(2007a).The Lucifer Effect: Understanding how good people turn evil.New York: Random House.

Zimbardo, P.(2007b).“The Stanford Prison Experiment: A simulation study of the psychology of imprisonment conducted at Stanford University.” Retrieved June 2, 2007, from http://www.prisonexp.org

Zimbardo, P.G., Maslach, C., & Haney, C.(1999).Reflections on the Stanford Prison Experiment: Genesis, transformation, consequences.In T.Blass (Ed.), Obedience to authority: Current perspectives on the Milgram paradigm (pp.193–237).Mahwah, NJ:

Erlbaum.

研究38 从众的力量

Asch, S.E.(1955).Opinions and social pressure.Scientific American，193, 31-35.

你觉得你是从众的人还是独立的人？我们大部分人更乐意别人认为自己是既从众又独立的，既不怪异、不叛逆，又有个性、有独立思考的能力。几十年以来，心理学家对从众这一概念始终非常感兴趣。如果你能想到心理学的研究不仅关注解释人类行为，而且更重要的是在于揭示其背后的原因，那么你也就理解为什么会这样了。人们愿意顺从他人的效应可以帮助我们理解人类行为的根源。

当心理学家谈到从众时，他们指的是当某一个体成为某个团体的成员时，其行为常追随团体的某种行为模式。通常，这种一个团体对其成员行为的不成文规定及约束被称为“社会规范”。认真思考一下，你或许会想起，在生活中有时你的行为与你的态度、信仰、道德并不同步或一致。很有可能，每当发生这种情况的时候，你正处于某一群体之中，群体中的每一个人都那样做，所以你也就跟随着那样做了。从众对于我们的行为而言是一种强大的力量，甚至可以说，这种力量在某些时候会使我们的行为偏离自身的内驱力，做出一些原本不可能做的事。所以，从众非常值得行为科学家投注兴趣并开展研究。然而，直到20世纪50年代初才有人对从众进行科学的研究，这个人就是所罗门·阿希（Solomon Asch）。他关于从众的系列实验为我们提供了有关从众行为的大量新信息，而且为未来的研究开辟了许多道路。

理论假设

假如，你与一群经常见到的人在一起，比如你的朋友或同事。他们正在讨论一些有争议的问题或政坛候选人。你很快发现，在这个群体中，其他人观点一致，而唯独你的观点与别人的恰恰相反。当他们询问你的意见时，你会怎么做呢？或是选择冒着被排斥的风险，不顾后果地坚持说出自己的真实想法；或是选择和群体保持一致，即使你并不同意这种观点；或者，如果有可能的话，你会完全回避这个问题。

阿希想要了解从众的需要对我们的行为到底有多大影响力。虽然从众经常会涉及一些笼统模糊的概念，如认同他人的态度、伦理、道德和信仰体系等，但是，阿希选择了集中研究一种更加显而易见的从众类型——知觉从众，即人们倾向于在多大程度上遵照他人关于世界的知觉（所看、所听、所尝、所闻和所触）。阿希通过一个简单的视觉比较任务来检验从众行为，这样他就可以在一种受控制的实验情境中研究这一

现象。

如果从众所具有的影响力真如阿希以及其他一些人所认为的那样强大的话，那么研究人员就可以通过让群体给个人施加压力来操纵一个人的行为，这就是阿希所从事的研究。他设计出一系列非常巧妙的实验，采用相似的方法来验证从众的力量。

方法

视觉材料由成对卡片组成，每对卡片的其中一张上画有三条不同长度的垂直线（称为比较线）；另外一张卡片上画有一条垂直线（称为标准线），其长度与前一张卡片中三条线中的一条等长（见图10—3）。接下来，便是实验的具体程序：假如你是自愿参加“视知觉研究”的一名被试，你到达了实验室，发现其他七名被试已经坐成一排。你坐在这一排的最后一个空位上。实验者呈现出一对卡片，要求你判断三条比较线中的哪一条与标准线的长度相等。你观察了一下卡片，并很快做出了正确的回答。从离你最远的那个被试开始，主试要求他们回答同样的问题。每个被试都给出了正确的答案，轮到你的时候，你也很清楚地给出了正确的答案。然后主试更换了卡片，又进行了同样的过程。毫无问题，你与其他成员一样，又一次给出了正确的答案。然而，在下一轮测试中，奇怪的事发生了——主试呈现卡片后，你立即想好了正确答案（毕竟，这并不困难），但是在其他被试回答这一问题时，他们却都选择了错误的答案，而且他们竟都选择了同样的错误线条。现在，轮到你作答了，你犹豫了，你简直不能相信眼前发生的一切。难道所有的人都瞎了眼吗？正确的答案是显而易见的，难道不是吗？你们到底是瞎了，还是疯了？现在你必须做一个决定。你是继续坚持你的观点（毕竟，正确的线条就在你的眼前），还是顺从其他人的意见？



图10—3 阿希线条判断任务卡片的一个例子

资料来源：引自P.32。

现在你可能已经意识到了，在实验室中的其他七个“被试”根本就不是真正的被试，而是实验者的助手。他们从开始就知晓实验内容，而且回答也是规定好的，当然，这也是从众研究的关键。那么，在这项研究中，真正的被试是如何回答的呢？

结果

每位被试测试数次，其中约有75%的人至少有一次与团体的不正确共识保持了一致。综合所有实验结果，被试服从于团体做出错误回答的次数约占1/3。为了确保线条的长度可以准确判断，主试要求控制组中的被试单独写下他们对线条比较问题的判断，这些被试的正确率为98%。

讨论与相关研究

在阿希的研究中，团体压力对从众行为的有力影响得到了清楚的体现。如果个体愿意服从于团体，那么他们对明显的错误有时也不具判断力，而这种影响力在现实生活中究竟有多大？团体在何时会发挥更强大的作用？问题在何种情况下会变得更模棱两可呢？从众效应是影响人类行为的一个重要因素，这个多年来人们一直反复思索的问题现在已得到了科学证实。

在心理学领域中，阿希的研究结果在两个方面显得极为重要和关键。首先，如前所述，他首次清楚而科学地证明了社会压力对从众行为的影响力；其次，或许也是更重要的，即这项早期研究引起了大量后续研究，其作用持续至今。自阿希的早期研究开始，该类研究的数量急剧增加，它们极大地丰富了我们对从众效应的了解，也使我们明白究竟是哪些因素决定了从众效应对人类行为的影响。下面是其中的一些研究发现。

1.社会支持。阿希在同样的实验上做了细微的变动，他对七名助手的回答进行改变，使其中一名助手在测试条件下给出正确的回答。在这种情况下，仅仅有5%的被试同意团体的一致性意见。很明显，你只需要一个同盟者就能“坚定立场”并抵抗从众效应的压力。这项研究发现还得到了后来几项研究的支持（例如，Morris & Miller，1975）。

2.团体的吸引力及成员的归属感。后来的一些研究证明，个体越为某一团体所吸引，对这一团体越有归属感，那么他就越有可能顺应该团体的态度与行为（参见Forsyth，1983）。如果你喜欢某一团体而且觉得自己是他们的一员（他们是你的参照群体），那么你顺应于该团体的倾向性将是非常强烈的。

3.团体的规模。首先，阿希和其他研究人员的研究都证明，从众的倾向性随团体规模的增加而提高。然而，进一步的研究发现，这种联系并非如此简单。确实，从众的倾向性随团体规模的增加而提高，但这仅限于人员数量为6到7人的团体。当其规模超出这个数字时，从众效应的水平不再增高，甚至在某种程度上有所降低。这

一现象在图10—4中以图解的方式呈现出来了。阿希认为，这可能是由于随着团体成员数量的增多，人们可能会开始怀疑其他成员是有目的地合起伙来影响他人的行为，作为回应，他们开始抵抗这种显而易见的压力。

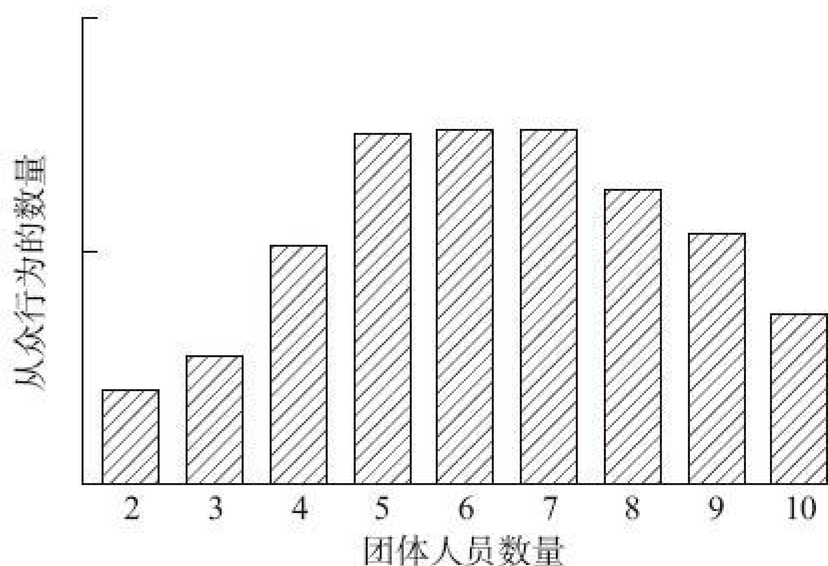


图10—4 团体规模与服从的关系

注：引自P.35数据。

4.性别。你认为男人和女人在从众的倾向性和自发性方面会有不同吗？继阿希的工作之后，有早期研究指出，女人似乎比男人更愿意从众。这种性别差异成为一种强有力的证据，频繁地出现在心理学文献之中。然而，近期的研究对这个结果提出了质疑。现在看来，许多早期研究（多由男性实施）无意中创造的一些实验条件，在当时对于男性比对于女性来说更熟悉、更舒适。心理学家明白，当人们处于那些不知如何表现才恰如其分的情境中，会表现出更多的从众倾向。因此，关于女性有较强从众倾向的结论也许只是一种系统误差，由方法上微小（且无意识）的偏差所致。在更好的控制条件下进行的一些研究并未发现从众行为有性别差异（关于性别问题的讨论参见Sistrunk & McDavid，1971）。

有关从众问题的研究还涉及其他众多领域，其中包括文化影响、在做出与从众相关的决定时的信息总量、社会规范、个人隐私，等等。

批评

阿希对从众效应的研究工作得到了广泛的支持和认可，此外该研究也在多种条件下被重复进行。对于这项研究的质疑则关注阿希的研究发现能否推广到实验室以外的现实生活中。换句话说，被试在实验室里对一些线条的回答与现实生活中的从众行为是否存在显著相关呢？这是对于所有在实验室控制条件下研究人类行为实验的效度的质疑。这种质疑的理由是：“也许被试可以在线段长度这种并不重要的琐事上与团体成员保持一致。但是，在现实生活中，特别是在重要事情上，他们不会那么轻易地表现出从众行为。”然而，虽然现实生活中的从众事件更有意义，但现实群体在从众事件上给予个体的压力也会相应增加。

近期应用

我们从以下这项研究中可以看出，阿希的工作对有关重大社会问题的研究依旧产生着重要影响。该文调查的是年轻人卷入不安全的性行为的原因（Cerwonka，Isbell & Hansen，2000）。大约有400名年龄在18~29岁之间的学生接受了不同方式的测验评估。测验内容包括AIDS/HIV高风险性行为知识（例如未使用安全套、多位性伙伴、酒精和其他药物滥用以及性经历等）。研究显示，众多因素能预测高风险性行为，其中包括基于同辈群体压力的从众。可以看出，对于从众如何影响人们选择性行为的理解可能成为我们对抗HIV持续传播的颇具价值的工具。

另一项引人关注的研究与阿希1955年发表的文章一脉相承，该研究考察为什么男性比女性更少寻求帮助，即使他们急需得到帮助（Mansfield等，2003）。这篇文章以下面这个（老）笑话引出：“为什么摩西用了40年时间在沙漠中徘徊，因为他不愿意问路（P.93）”。这个笑话之所以有趣，是因为它触及有关男性和寻求帮助行为的刻板印象。当然，不问路通常不会造成严重的问题，但是男性也倾向于拒绝寻求医疗和心理健康救助，这可能是危险的，甚至是致命的。作者指出，阻止男性寻求帮助的一个主要因素就是从众。“在需要帮助的情境中，男性可能会不愿意寻求帮助，如果他们认为这样做会受到羞辱……如果一个男性在生活中非常钦佩对寻求帮助持不赞成或者贬低观点的人，那么他自己就不太可能寻求帮助”（P.101）。

最后提到的一点是，文化在从众中起着特殊的重要作用（Bond & Smith，1996）。在具有集体主义文化的国家，比如日本、印度，所进行的从众研究发现，人们的从众行为都远远多于具有个人主义文化的国家的人民，如美国（参见研究28中川迪斯对集体主义和个人主义文化的研究）。这些发现再次证明心理学研究决不能忽视文化对人类行为的实际影响。

Bond, R., & Smith, P.(1996).Culture and conformity: A preferencemeta-analysis of studies using Ash’s line- judgment task.Psychological Bulletin，119 (1)，111-137.

Cerwonka, E., Isbell, T., & Hansen, C.(2000).Psychosocial factors as predictors of unsafe sexual practices among young adults.Aids Education and Prevention，12(2)，141-153.

Forsyth, D.(1983).An introduction to group dynamics.Pacific Grove, CA: Brooks/Cole.

Mansfield, A., Addis, M., & Mahalik, J., (2003).Why won't he go to the doctor? The psychology of men's help-seeking.International Journal of Men's Health, 2, 93-109.

Morris, W., & Miller, R.(1975).The effects of consensus-breaking and consensus-preempting partners on reduction in conformity.Journal of Experimental Social Psychology，11, 215-223.

Sistrunk, F., & McDavid, J.(1971).Sex variable in conforming behavior.Journal of Personality and Social Psychology，17, 200-207.

研究39 你会伸出援手吗

Darley, J.M., & Latané, B.(1968).Bystander intervention in emergencies: Diffusion of responsibility.Journal of Personality and social Psychology, 8, 377-383.

在心理学研究的历史上，最有影响的事件之一并不是某项研究，而是由美国新闻媒体所报道的一则发生在纽约市的暴力悲剧事件。1964年的一天，一位年轻女士凯蒂·吉诺维斯（Kitty Genovese）在她经营的曼哈顿酒吧营业结束后返回公寓，她的公寓在皇后大街的一个安静的中产阶层居住区内。当她下车朝公寓方向走去的时候，她遭到一个持刀男人的恶意袭击。那男人刺了她好几刀，她大声喊救命。一个邻居从窗口大声警告那男人："放开这个女孩"，歹徒正欲图逃走，但后来他又返回来将吉诺维斯击倒在地，并继续刺杀她。袭击仍在继续，女孩一直呼救直到最后有人报警。警察接到报警后两分钟便赶到了现场，但吉诺维斯当时已经死了，袭击者也不知去向。袭击行为持续了35分钟。警察在调查这一事件中发现，公寓周围共有38个人目睹了这一袭击事件，但最终只有一人报了警。一对夫妇（他们说他们以为已经有人报了警）把两把椅子移到窗前去观看这一暴力事件。杀害吉诺维斯的人温斯顿·莫斯利（Winston Moseley），现年将近70岁，至今被关押在位于纽约北部戒备最森严的监狱里。在他47年的牢狱生涯中，他提请的14次假释都被否决了。在最近的2011年假释听证上，他的假释再一次被否决，直到2013年他才能再次提请假释。

如果有一个人能早些向吉诺维斯伸出援手的话，她也许就能幸免于难。这看上去似乎是邻居们漠不关心，没有试着阻止这场暴力事件，这令纽约市民和全体美国人感到震惊。人们想知道这其中的原因，同时他们谴责居住在大城市里的人们情感淡漠，谴责皇后大街的居民们，谴责人类的本性。

吉诺维斯的悲剧激发了心理学家的研究兴趣。作为科学家，他们不打算谴责谁，而是试图理解阻止所有的人对受害人伸出援助之手的心理因素。在心理学研究领域中有一个概念，行为科学家称之为"亲社会行为"，或者说是产生积极社会后果的行为。这个研究领域的内容包括利他、合作、抵御诱惑以及帮助行为。如果你目睹了某人需要帮助的紧急情境，有许多因素会影响你做出挺身而出并提供帮助的决定。纽约大学的约翰·达利（John Darley）和哥伦比亚的比勃·拉特内（Bibb Latané）就是两位研究这些影响因素的社会心理学家。他们把这种在突发事件中的帮助他人行为称为"旁观者干预"行为（而在吉诺维斯的案件则叫未干预行为）。

你曾经遭遇过真实的突发事件吗？与你在电视中看到的、报纸中读

到的相反，突发事件并不常见。达利和拉特内估计，平均每个人一生中遇到的突发事件不多于六件。这既是好事，也不是好事，好的原因显而易见，不好的原因是因为当你发现自己正面对一个突发事件，而你又必须做出决定时，你却没有任何可利用的经验。社会要求我们在遇到突发事件时要伸出援手，但通常的情况就像著名的吉诺维斯事件那样，我们没有这么去做。是因为我们经历那样的事件太少，以致使我们只是不知道该做什么？是因为居住在城市里的人感情冷漠吗？或者是因为人的本性本身就是冷漠的？

自吉诺维斯谋杀事件发生之后，达利和拉特内分析了旁观者的反应。他们提出的理论是：正是由于观看事件的目击证人太多而降低了任何一个个体提供帮助的意愿。他们决定用实验的方法来验证这一理论。

理论假设

常识可能会告诉你，在一件突发事件发生时，在场的旁观者数量越多，他们干预的可能性就越大。达利和拉特内却提出了相反的假设：他们认为没有人前去帮助凯蒂·吉诺维斯的原因是存在一种被他们称为“责任扩散”的现象。也就是说，在突发事件中，旁观者越多，人们心里的某种想法就越强烈，即“有人会去帮助他（或她）的，我就不必去了。”你是否曾在繁忙的街道上看见过车祸，或者当车祸发生不久你恰好到达了出事地点？你很可能驾车驶离现场，并认为现在一定已经有人叫了警察或救护车，所以你不觉得自己有责任也这样做。但是，想象一下，假设你在荒郊野外，周围没有其他人，你遇到了同样的事件，此时你会有不同的反应吗？我们中的大多数人会回答“是的”。

责任扩散的概念是本项研究的理论基础。研究的挑战在于在一个可控的实验室情境中重新创设类似吉诺维斯事件情境，这样就可以操纵和系统检验各种变量。达利和拉特内精心设计了一系列实验来验证其理论观点。

方法

很显然，研究者不可能为了进行实验而重新创设一种与凯蒂·吉诺维斯谋杀案完全相同的情境。因此他们必须设计另一种与真实的突发事件非常相似的情境以便观察旁观者的干预行为。在这个实验中，达利和拉特内告诉纽约大学选修心理学课程的学生，他们两人对一些与大学生有关的话题非常感兴趣，想了解大学生是如何在激烈的竞争中适应大学生活、城市环境以及他们正面临着什么样的个人问题。他们要求学生坦诚地与其他学生讨论自己的问题，但为了避免难堪和不愉快情况的发生，学生们单独待在相互隔离的房间里，彼此之间通过内部通信联络系统进

行交谈。研究者还告诉他们，这种内部装置每次仅允许一名学生讲话。每位学生有两分钟的讲话时间，之后下一位同学的麦克风就会开启。

所有这些隐藏了真实实验目的的设计，是为了能够从被试那里获得真实、自然的行为反应。这其中最主要的方法就是把这些参与研究的学生分在三种不同的实验条件下。第一组的被试相信他们仅仅能与另外一个人交谈；第二组的被试相信他们可以通过内部通信系统与另外两个人交谈；第三组的被试被告知，他们可以与线上的另外五个人交谈。事实上，每个被试都是独自一个人，他们从“内部通话系统”中听到的其他的声音都是事先准备好的录音。

现在，实验组的人数已安排好了，接下来就是要设计一些突发事件。研究人员决定非常逼真地设计一个被人们普遍认为是突发事件的癫痫发作场景。当被试们与其他“学生”在内部通信系统开始讲话时，被试听到的第一个说话者是男性，他提到他的困难主要集中在学习和适应纽约的生活氛围方面，然后他又尴尬地补充说，自己有时候也会发作严重的癫痫，特别是在压力过大的时候。然后，便轮到下一个学生讲话。在第一组中，被试听完第一个学生的谈话后，就立即开始讲话。而在其他两种条件下，被试在讲话之前会听到一个或多个其他学生的说话声。被试说完后，又轮到第一个学生。这时候突发事件发生了。第一个学生一开始声音很正常，不久癫痫开始发作（记住，这完全是录音机里发出的声音）。达利和拉特内在后来的报告中引用了“癫痫发作”的全部内容：

“我、我……想，我需要……有……如果有人……能给我一点给我一点帮助的话……因为我……现在真的有麻烦，如……果有人能帮助……我的话，那……太感激……了。我的癫痫就要发作……我需……要一些……帮助，如果有人愿意帮助我……（哽咽声）我要死了，帮……助癫痫……（哽咽，然后就无声了）。”（pp.95–96）

对于被试来说，这很显然是一个突发事件。毫无疑问，那名“学生”正处在危难之中，需要立即得到帮助。为了分析被试的反应，达利和拉特内测量了每一种实验条件下被试帮助危难学生的百分率（帮助行为指的是离开小房间，通知实验者有学生癫痫发作）。他们还测量了被试对突发事件的反应时间和提供帮助所花的时间。研究者给被试4分钟的时间作反应，之后实验停止，研究者询问被试相关情况。

结果

这项研究的结果对研究者的假设提供了强有力的支持。随着被试认为共同在场人数数量的增加，能够迅速向研究者报告癫痫发作的人数百分比大大减少（见图10—5）。而在所有最终提供帮助的人中间，时间拖延总量也随旁观者人数的增加而增加。第一组被试平均的反应时间不到一分钟，而第三组被试的平均反应时间却超过了三分钟。报告突发事件的被试总人数在各组间也有显著的变化。在四分钟内，第一组被试百

分之百地报告了突发事件，第二组被试中只有85%的人报告，而第三组被试中只有60%的人作了报告。

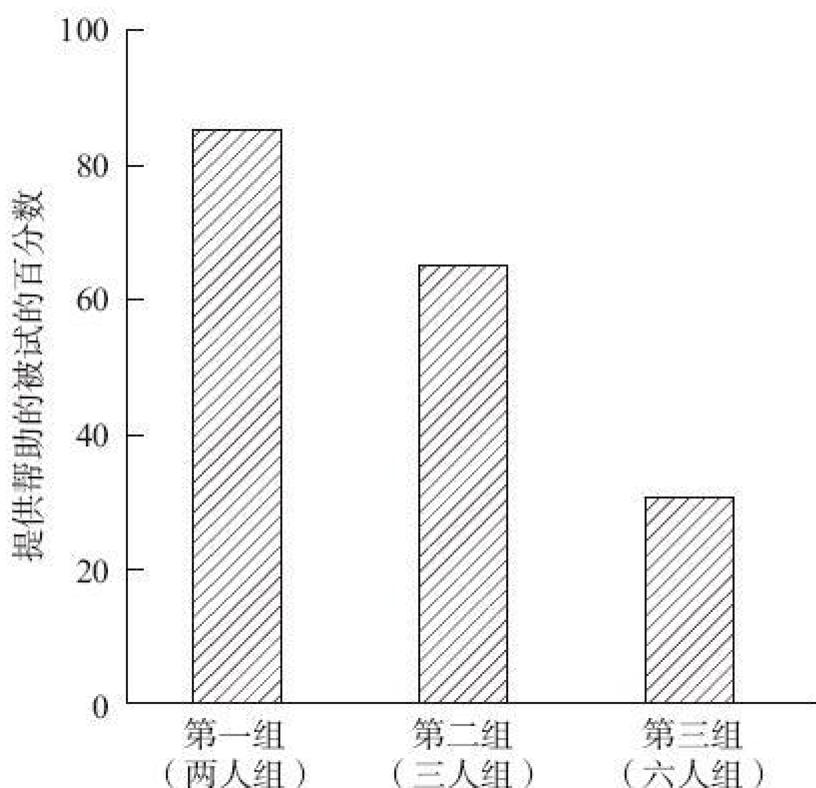


图10—5 在当事人癫痫发作时，各种实验条件下迅速提供帮助的被试人数

资料来源：引自P.380数据。

讨论

就像凯蒂·吉诺维斯谋杀案这一现实生活事件一样，你也许会认为这项研究中的被试对癫痫发作者非常冷漠、令人心寒。然而达利和拉特内很快指出，第二组和第三组被试（或吉诺维斯的邻居们）实际上并不是这样的。所有的被试都报告了在事件发生时，他们感到极度的焦虑和不适，还有人出现了紧张的躯体症状（如手发抖、掌心出汗）。研究者得出的结论是，被试行为结果的不同一定与被试认为的在场人数的不同有关。别人的在场改变了你的行为，这就是大家所知道的“社会影响”心理学原理的作用。很显然，社会影响在本研究中扮演着一个重要的角色。但是，我们仍然有疑惑，为什么其他人的在场会对我们有如此巨大的影响？

达利和拉特内声称已有证据支持他们的“责任扩散”理论。随着团体内人数的增多，被试在突发事件中采取行动的个人责任就减少了。对于第二组和第三组的被试来说，他们很可能会认为有人已经解决了这个问题。此外，当其他人在场的时候，人们不仅仅感觉到责任的分散，而且会觉得如果不提供帮助所带来的内疚和良心责备也更少一些。在我们的文化中，帮助别人是积极的行为，拒绝或不肯帮助别人会带来羞愧感。如果在突发事件中只有你一个人在场，那么你不提供帮助所要承担的后果要比有其他人在场时大得多。

对这类社会影响的另一种解释是心理学家称之为“评价恐惧”的现象。达利和拉特内认为，当别人在场时，我们没有去帮助当事人的部分原因是我们害怕难堪或被嘲笑。设想一下，如果你有帮助别人的愿望，而别人并不需要或不想让你帮助时，你将感到多么尴尬。我记得在我年轻的时候，有一次我与一帮朋友在邻居家的游泳池内游泳，当我正准备从跳板上跳水时，我看见邻居13岁的女儿正躺在泳池的底部，我看了一下四周，似乎没有一个人意识到这个显而易见的紧急状况。她是溺水了还是在开玩笑？我无法确定。正当我准备大声呼喊并跳下去营救时，她懒洋洋地游出了水面。因为担心自己出错、害怕过度反应的尴尬，我犹豫了整整30秒。我们许多人都经历过这类事情。问题是这些经历教给我们的是一种错误的东西：帮助他人的行为很可能使你看上去很愚蠢。

研究发现的意义

由于本项研究以及一些其他研究，达利和拉特内在助人行为和旁观者干预领域内成了开创性的研究者。他们早期的大量研究收录在他们的著作《反应冷漠的旁观者——他为何不帮助？》（The Unresponsive Bystander：Why Doesn’t He Help）一书中（Latané和Darley，1970）。在该书中，他们所描绘的助人行为模型在被有关助人研究的心理学文献广泛认可。他们提出大多数人介入某一突发事件前一般要经历5个步骤。

1.你——一个潜在的帮助者——必须首先意识到某一紧急事件正在发生。在本章的研究中，紧急情况的产生是毫无疑问的。但在真实的生活中，你也许行色匆匆或将注意力集中在另一件事上，可能完全没有注意到紧急情况的发生。

2.你必须对情境作出判断，即某个人是否需要帮助。在这一点上，害怕难堪的心理发挥着作用。在本研究中，情境是清晰的，需要帮助也是非常明确的。然而，在现实中，大部分潜在的突发事件具有某种程度的不确定性和模糊性，如我在游泳池游泳的那个例子。或者想象一下，你看见一个人正步履蹒跚地穿过繁忙的人行道，你会想他是病了，心脏病发作，还是喝醉了？你对情境的判断会影响你对是否介入此事的决定。在吉诺维斯事件中，许多没有提

供帮助的人声称，他们都以为这是情侣间的争吵而不想介入其中。

3.你必须承担个人的责任。如果在突发事件中，你是唯一的旁观者，那么你会立即承担这种责任。然而，如果有其他人在场，你也许会让他们替你承担责任。这一步是本项实验所关注的重点。在突发事件中，在场的人越多，责任扩散就越严重，也就越有可能发生无人提供帮助的情况。

4.如果你承担了责任，那么你必须决定你该采取什么行动。如果你不知道该做什么或者觉得自己没有采取适当行动的能力，你将不会提供帮助。在达利和拉特内的研究中，被试只需要向研究者报告有学生癫痫发作，完全不涉及能力问题。但如果有一群人看见一辆车撞倒了一个行人，这群人中有医生、护士或是其他护理人员，那么他（或她）比其他人更有可能介入其中，因为他们会更知道自己该做些什么。

5.当你决定该采取什么行动之后，你必须去做。仅仅知道自己该做什么并不能保证你就一定会去做。现在，你必须掂量帮助别人的利与弊。你愿意独自介入有一人持刀或两个人都持刀的打斗中吗？对于车祸中的受害者，你是否能帮助他们还是可能会把事情弄得更糟呢（又是一个能力的问题）？如果你介入了，你会被起诉吗？如果你本想努力去帮助别人，但结果却像个傻子，又将如何？许多问题取决于情境，在你采取真正的行动以前，这些问题会在你的脑海里一一闪现。图10—6说明了助人行为是如何被中断的或在某一阶段被阻断的。

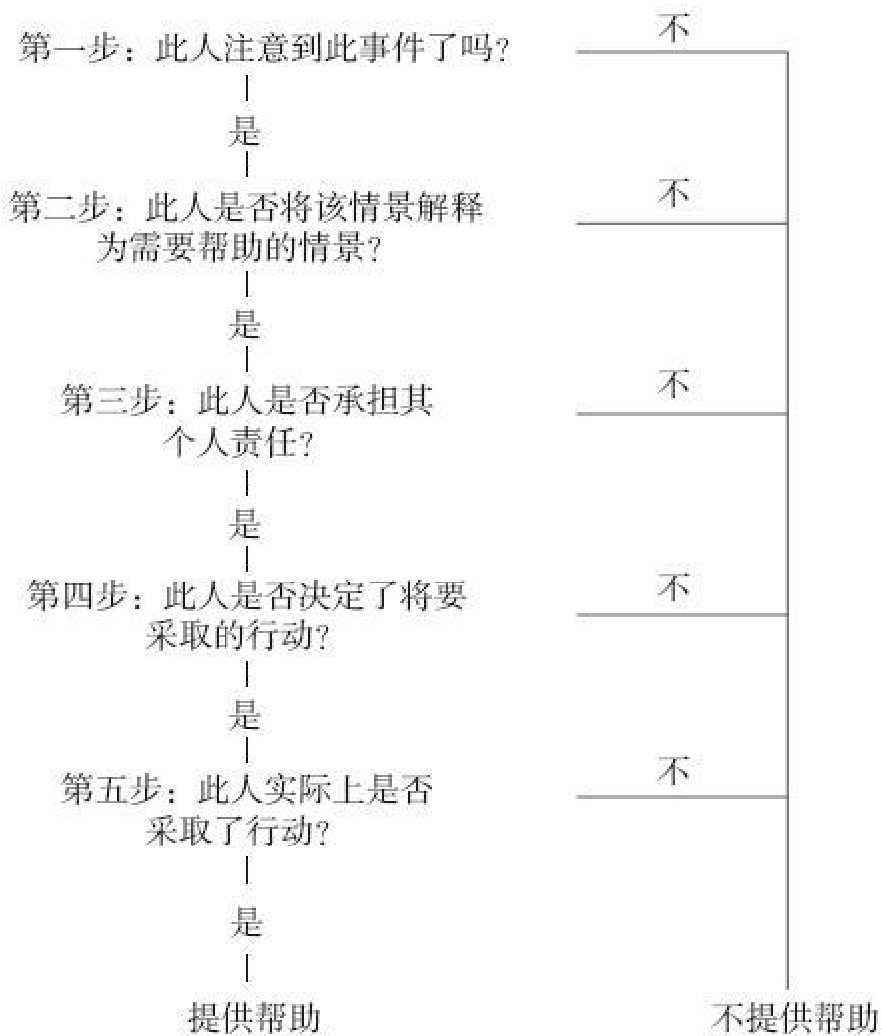


图10—6 拉特内和达利的助人模型

后续发现与近期应用

在凯蒂·吉诺维斯谋杀案和我们在这里讨论的这个实验中，所涉及的旁观者都是彼此无法交流的。如果这些旁观者彼此能相互看见，并且可以相互交流的话，你认为后果会怎样？当他们可能被其他人评价时，是否会更可能介入其中？达利和拉特内认为，在某种情况下，即便是关系很近的群体也不比个人更乐意提供帮助。他们的理论认为，当突发事件在某种程度上比较模糊时，这种情况就更有可能发生。

举例来说，设想一下，你正坐在候车室里，然后通风口开始冒烟。

你忧心忡忡地看了一下房间里的其他人。他们显得出奇地平静，对此毫不在意。你会觉得自己对冒烟的反应一定有些夸张，所以决定不采取任何行动。因为你若采取了行动，而却错了（也许它不是烟，而是从隔壁房间里冒出来的水蒸汽或其他什么东西），你将会感到羞愧和尴尬。然而，你并不知道，房间里的每一个人其实与你有同样的感受，并且他们也试图把这种感受隐藏起来，像你一样在避免难堪！而对于冒烟，却没有一个人采取任何行动。这听起来有些不可信，对吗？但事实确实如此。

拉特内和达利（1968）在随后的一项研究中设计了一个类似上述的情境来验证这一观点。一些心理学系的学生自愿来参加一项访谈，他们被告知将共同“讨论在城市大学中的一些生活问题”。他们为接受访谈来到预先约定的地点，并被安排在一个房间里等候，填写一份预备问卷。几分钟后，烟雾开始从某个通风口进入房间，研究中使用的烟雾是一种特殊的化学混合物，对被试不造成任何危险。不久，烟变得越来越浓，以致于房间里的视线开始变得模糊。主试开始计时，看被试会过多长时间才把冒烟的情况报告给主试。一些被试被安排在单独的房间里；还有一些被试与另外两个人（主试的助手）在一起，被试认为这两个人也是被试，当烟出现时，这两个人表现得若无其事；其余的被试则三人一组待在房间里。拉特内和达利的研究结果又一次支持了他们的理论：在单独组中，有55%的被试在两分钟之内报告了冒烟的情况；其余两个组仅有12%的被试报告了情况；另外，四分钟之后，单独组有75%的被试开始行动，而在其他组中，仍然没有被试报告冒烟一事。

关于害怕尴尬让人们对帮助他人犹豫不决的更多证据来自另外一项研究，在这项研究中综合考察了害羞人格测量、对负面评价的恐惧（FNE），以及被试帮助他人的意愿（Karakashian et al.，2006）。被试先填写测量害羞和恐惧负面评价的问卷。然后，他们会面临着帮助一位女性同伴的机会，有些被试是与这位女士单独在一个房间，另一些被试则除了这位女士以外还有另外两个同伴（研究助手）在一起。根据达利和拉特内的研究结果，无论被试人格测量的分数如何，与没有其他旁观者在场的条件相比，有两位旁观者的情境中被试的帮助行为会显著地减少。除此以外，在无旁观者条件下，那些在恐惧负面评价和害羞测量上得分高的被试比其他被试更少地提供帮助；而在有两位旁观者在场的条件下，不同被试提供帮助行为的可能性相同。这可能看起来与你的直觉相悖。通常人们直觉认为害怕被负面评价的或者是害羞的人在有其他人在场的时候可能更少会提供帮助，是吧？事实上并不完全如此。让我们这么考虑：当其他人在场的时候，一个害羞的人觉得提供帮助的压力更小（由于责任扩散），所以他（或她）其实有“理由”像其他旁观者一样避免提供帮助；另一方面，如果没有其他旁观者在场，对于（潜在的）负面评价的恐惧就生效了，害羞的人比不害羞的人提供帮助的可能性就更少。研究者对于此项研究的说明如下：

由于在社会条件下（有其他人在场）存在责任扩散，被试面临较少的是否提供帮助的决策。这时，由于提供帮助的想法本身不多，也就不用担心受到糟糕的评价，那

么恐惧负面评价就不成问题了。而在非社会条件下（没有其他旁观者），被试是单独一个人，具有提供帮助的全部责任，这样他就必须要决定是否采取行动（Karakashian et al.,2006，P.30）。

另一项研究揭示了在我们的想象中（而不是现实生活中）旁观者效应和责任扩散的力量。包括达利在内的一个研究小组进行了一项题为《群体思维：固有的旁观者效应》的研究，结果发现，仅仅是设想自己是群体的一员就会改变帮助行为（Garcia et al.，2002）。在这项研究中，要求被试将自己设想为群体的一员，或者是与另外一个人在一起。然后要求所有的被试为一个慈善机构捐款。想象自己是群体一员的个体比想象和另外一个人在一起的被试捐出更少的钱，而且感觉自己的个人责任更少。这些结果显示，当我们作为群体一员的时候，我们的大脑会立即把握机会，认为自己的个人责任变少。

结论

虽然，这项研究的结果看起来确实让我们对帮助他人的内在倾向有些悲观，但是你应该认识到，这些所涉及的都是人们未能伸出援手的一些极端特殊情况。生活中每天都会发生人们去帮助他人、利他甚至是英雄主义的行为。这项研究之所以重要，是因为它不仅解释了使人困惑的人类行为，而且还有助于改变这种行为。当人们更多地认识到旁观者效应的时候（正如你现在已经知道其中原理），即使有其他人在场，他们也会作出更多的努力去干预危急情况。事实上，有研究已经证明，当人们了解了旁观者效应后，他们会更乐意向危难中的人们伸出援手（Beaman et al.，1978）。关键在于，危急时刻，千万不要以为已经有人提供了帮助或者一定会有人提供帮助。我们应该永远像自己是唯一的旁观者一样去行动。

Beaman, A., Barnes, P., Klentz, B., & Mcquirk, B.(1978).Increasing helping rates through information dissemination: Teaching pays.Personality and Social Psychology Bulletin, 4, 406-411.

Garcia, S., Weaver, K., Darley, J., & Moskowitz, G.(2002).Crowded minds: The implicit bystander effect.Journal of Personality and Social Psychology, 83, 843-853.

Karakashian, L., Walter, M., Christopher, A., & Lucas, T.(2006).Fear of negative evaluation affects helping behavior: The bystander effect revisited.North American Journal of Psychology, 8(1)，13–32.

Latané, B., & Darley, J.M.(1968).Group inhibition of bystander intervention in emergencies.Journal of Personality and Social Psychology，10, 215-221.

Latané, B., & Darley, J.M.(1970).The unresponsive bystanders: Why doesn’t he help? New York: Appleton-Century-Crofts.

研究40 无条件服从

Milgram, S.(1963).Behavioral study of obedience.Journal of Abnormal and Social Psychology, 67, 371-378.

如果一个在你之上的权威人士命令你对另一个人施加电压为350伏电压的电击，原因仅是因为这个人在回答多项选择问题时答错了，你会听从命令吗？没有人会这样做。如果有人愿意去做这件事，你会认为他或她是一个残酷的人或是虐待狂。耶鲁大学的斯坦利·米尔格拉姆（Stanley Milgram）所做的这项研究，试图检验服从权威的想法，并得出了一些令人深感不安的结果。

米尔格拉姆的服从研究和津巴多的监狱研究（见研究37）在整个心理学史上极负盛名，受到人们的广泛认可。几乎在每一本普通心理学和社会心理学的教科书中都会提到该项研究。如果你与心理学专业的学生交谈，他们中的大多数人对这些研究要比对其他研究更熟悉。由这项研究引发米尔格拉姆写了一本关于服从心理的书，拍了一部与此相关的专题片，该片在大学心理学课上反复放映。这个实验不仅涉及有关服从行为的讨论，而且还影响了关于在心理学研究中采用人类被试的伦理问题的全部讨论。

米尔格拉姆的这项研究方案来自于他的一个想法，即他想用科学的方法调查人们怎么会仅仅因为别人的命令而对他人施加巨大的伤害。米尔格拉姆特别提到纳粹在二战期间犯下的可怕暴行，也在更广泛的意义上提及此前及现在因服从命令而发生的不人道行为。米尔格拉姆认为，在某些情况下，人类服从倾向是如此强烈而根深蒂固，它抵消了个人的道德、伦理，甚至是同情。

在行为科学家决定要研究某些人类行为复杂面的时候，他们的第一步就是要对发生行为的情境进行有效控制，以便能用科学的方法对它进行研究。这对于研究者而言是一个极大的挑战，因为要在实验环境中模拟现实生活是一件非常困难的事情。米尔格拉姆的问题是如何创设一个受控制的情境，在这个情境中一个人可以命令另一个人去体罚第三个人，且实际上并没有人真的受到伤害。这就是对研究者的一个挑战！

理论假设

米尔格拉姆最初的理论假设是，人类有一种服从权威命令的倾向性，即使这个命令违背他们自己的道德和伦理行为准则。例如，米尔格

拉姆认为，虽然许多人从未有意伤害过别人，但在他们认为是权威人士的命令下，他们很有可能会对受害人做出伤害行为。

方法

也许这项研究最巧妙的部分就是米尔格拉姆开发的在实验室里验证服从力量的技术方法。米尔格拉姆设计了一个看起来非常骇人的电击装置，一个电子设备有30个调节开关，每个开关上贴着标签表示不同的电压，电压从30伏开始，每次以15伏为单位递增，一直增加到450伏。这些开关被分成三组，分别标着“轻微电击”、“中等电击”和“高压危险电击”。主试欲用这个装置让权威人士命令被试不断地增加电压去电击另外一个人。在你得出米尔格拉姆是一个施虐狂的结论前，请注意，那是一个非常逼真的模拟电击装置，它并不会让人真的受到任何痛苦的电击。

这项研究中的被试是40名年龄在20～50岁之间的男士，其中15人是熟练工人或不熟练工人，16人是售货员或商人，9人是专业人员。研究者通过报纸广告或直接发信的方式征集志愿者，邀请他们参加耶鲁大学的一项关于学习与记忆的有报酬研究。每位被试都是单独参加实验。为了获得较多的被试，研究者付给每位被试4.5美元（记住，这是1963年的美元，相当于现在的30美元），并清楚地告知所有被试，报酬在他们来到实验室后便付清，而且不管在他们到来之后发生什么事情，这些报酬都是他们的。这就确保被试知道他们可以随时退出研究，而且不会因担心得不到报酬而感觉被强迫以某种方式行事。

除了被试之外，研究中还有两位关键人物：一位是研究者的助手（一名47岁的会计），扮演另外一名被试；另一位是演员（穿着一件灰色的实验室工作服，看上去非常正式），扮演主试的角色。

当被试到达耶鲁大学社会互动实验室时，工作人员安排他坐在另一名被试（研究者的助手）的旁边。很明显，实验的真正目的是不能告诉被试的，因为那样将会彻底改变他们的行为。所以，实验者给被试讲了一个假故事，解释说这是一项对“学习中的惩罚”效应的研究。然后让被试通过从帽子里抽签来决定谁当老师，谁当学生。这个抽签是事先安排好的，因此真正的被试总是抽到老师，而研究者的助手则总是抽到学生。请记住，那名“学生”以及“主试”都是研究者的同伙。

然后，“学生”被带到隔壁房间，被试看到工作人员将他绑在椅子上，缠上电极（加入电极胶以“避免皮肤起泡或烧伤”），这些电极和邻近房间里的电击装置连在一起。虽然学生的手臂被绑着，但是他仍可以够得着标有A、B、C、D符号的按钮，并以此来回答隔壁房间里教师提出的问题。

主试向“老师”和“学生”清楚明了地交代了学习任务。简单地说，该

任务是让学生对各种各样的单词配对进行联想记忆。单词表很长，因此这是一个并不简单的记忆任务，扮演老师角色的被试把每对单词读给学生听，然后检查学生的记忆情况。主试要求这些被试在学生作出错误的反应时给予电击惩罚。最重要的是，每增加一次错误反应，“老师”就要将电压强度提高一级。所有的这一切伪装得极为逼真，没有一名被试对电击的真假表示怀疑。

扮演学生的助手的反应是预先安排好的，所有的被试都以相同的顺序作正确或错误反应。而且，当电击总量随着错误的反应增加时，“学生”开始从另外一个房间发出痛苦的叫喊（这也是预先安排好的，包括其叫喊的话语内容及其心脏不舒服的事实）。当电压达到300伏时，“学生”会猛撞墙壁，要求工作人员放他出去。超过300伏电压之后，他变得完全沉默，拒绝回答任何问题。主试告诉“老师”，用对待错误反应的方式来处理不反应的情况并继续进行应有的电击程序。

大部分被试在电压达到某一点时会转向主试询问是否继续进行电击。当这种情况发生时，主试会命令被试继续，且在此过程中很有必要向被试发出一系列语气渐重的命令。

命令1：请继续。

命令2：实验需要你继续。

命令3：继续进行是绝对必要的。

命令4：你别无选择，你必须继续。

通过记录每位被试拒绝继续进行的电压水平，就可以对服从行为进行测量。在这个电击装置上有30个开关，每位被试的得分在0～30之间。进行完全部过程，获得满分的被试被称为“服从的被试”，而在较低电压水平就中断的被试被称为“对抗的被试”。

结果

被试会服从主试的命令吗？他们最多会用多高的电压？你预测的结果是什么？想想你自己、你的朋友和一些普通人。你认为有多少人会使用所有30个电压水平，达到最高的450伏——高压危险电击呢？在讨论这个实验的真正结果之前，米尔格拉姆让一组耶鲁大学心理学专业的四年级学生以及许多其他同事预测结果。人们的估计最低是0%，最高是3%，平均为1.2%。也就是说，一百个人中会对“学生”施加最高电压电击的不足三人。

表10—2展现了令人惊讶的实验结果。在主试的命令下，几乎所有被试都将电压提升到了300伏的水平，直到“学生”猛击墙壁，要求离开

实验室，并拒绝回答问题为止。但是，最令人吃惊的是使用全部30个电压水平并使电压达到最大值的被试数量。

尽管有14名被试不服从命令，在达到最高电压之前中断了实验程序。但40名被试当中仍有26个，也就是65%的被试按照主试的命令继续进行实验，并使用了最高电压。这并不是说被试能心平气和或很高兴地去做这些事情，许多人表现出了极大的心理压力，并为受电击者的境况而担忧，甚至对主试非常愤怒。然而，他们还是服从了命令。

表10—2 被试执行的电击水平

电压等级数（伏特）	拒绝在该电压等级上继续执行电击命令的人数
轻微的电击	
15	0
30	0
45	0
60	0
中等的电击	
75	0
90	0
105	0
120	0
较强的电击	
135	0
150	0
165	0
180	0
很强的电击	
195	0
210	0

续前表

电压等级数（伏特）	拒绝在该电压等级上继续执行电击命令的人数
225	0
240	0
猛烈的电击	
255	0
270	0
285	0
300	5
极为猛烈的电击	
315	4
330	2
345	1
360	1
高压危险电击	
375	1
390	0
405	0
420	0
xxx---	
435	0
450	26

（引自米尔格拉姆，1964年，P.376）

研究者担心一些对他人施加电击酷刑的被试可能在心理上也遭受痛苦，尤其是在最后三次电击中，“学生”不再作出任何反应。为了帮助他们缓解这种焦虑，在被试完成实验后，研究者将这个实验的真正目的以及全部的实验程序，包括对他们的欺骗手段原原本本地告诉了他们。此外，研究者还对被试进行了访谈，了解他们在实验过程中的感受和想法。作为助手的“学生”也来与每位被试见了面。

讨论

米尔格拉姆对研究结果的讨论主要集中在两点上：首先是被试惊人的服从倾向。这些被试是同意参加一个关于学习的实验研究的普通人，绝不是冷酷的虐待狂。米尔格拉姆指出，从童年开始，他们就知道昧着良心去伤害别人是不道德的。那么他们为什么要那样做呢？主试是一个处在权威地位的人，但是你仔细想想，他到底有多少权力呢？他没有任何权力发号施令，被试若拒绝执行命令也不会有任何损失。很显然，是情境本身有一种力量，在一定程度上创设了服从的氛围。

第二个值得注意的关键是，在整个实验过程中，被试因执行了主试的命令而表现出极度的紧张和焦虑。研究者猜测，被试只要拒绝继续进行实验，这种不适感就会减轻。然而，这种情况并没有发生。米尔格拉姆引用了一个观察者的话（此人通过单向玻璃观察被试）：

> 我看到一位成熟稳重的商人进入了实验室，他面带微笑，充满自信。但在20分钟内，他就变得痛苦不堪，说话结结巴巴，很快就近乎于一种神经崩溃的状态……有时候，他用自己拳头敲着自己的前额，喃喃自语道："噢，上帝！让我们停止吧！"然而，他还是按主试的要求继续进行，直到进行完所有电压等级。"（P.377）

在文章的结尾，米尔格拉姆列举了几点理由来解释为什么这种特殊的情境会产生如此强烈的服从倾向。从被试的角度来看，原因主要有以下几点：（a）如果这项研究是由耶鲁大学发起的，那它一定是好事情，没有谁会怀疑这样的著名学校；（b）实验的目的似乎很重要，因为我是志愿者，所以我会尽力完成我的任务来配合研究者实现这些目标；（c）毕竟，"学生"也是自愿来这儿的，他对这项工作也有责任；（d）嘿，我是老师，他是学生，这纯属巧合——我们是抽签决定的。其实另一种情况也很有可能出现，即我是学生，他是老师；（e）他们为这事给我报酬，我要尽力做好；（f）我完全不知道心理学家以及被试的权力，所以我将屈从于他们的安排；（g）他们告诉我们，电击是痛苦的，但没有危险。

研究发现的意义

自从这项研究发表后的40多年以来，米尔格拉姆的研究仍占据着很重要的位置。米尔格拉姆本人用耶鲁大学以外的其他类似被试重复了这一实验过程，他们是义务参加实验的大学生志愿者和一些女性被试，每次得出的结果都是相似的。

另外，通过实施一系列的相关实验设计，米尔格拉姆进一步拓展了他的研究发现以揭示促进服从行为或限制服从行为的原因（参见Milgram，1974）。他发现，受害者（指"学生"）与"老师"的身体和情感距离的改变制约着服从倾向的强烈程度。当"学生"在另外一个房间，他与"老师"在彼此不见面也听不见对方的声音的情况下，出现了最高的服从率（93%的被试用了最高电压）。当"学生"与被试同在一个房间内，主试要求被试强迫"学生"把手放在电击板上，此时的服从率最低，

只有30%。

米尔格拉姆还发现，权威人物（指主试）与被试的身体距离也会影响服从行为。被试离主试越近，其服从倾向越强烈。在一种实验条件下，主试离开房间，并用电话对被试发出指令，结果被试的服从率仅为21%。

最后，有一个现象较为令人乐观，即当主试允许被试用他们想用的电压水平惩罚“学生”时，没有一名被试使用的电压超过45伏。

批评

米尔格拉姆的研究不仅影响着我们对服从行为的理解，而且使人们对用人类作被试的道德问题进行深刻的思考。即使没有一个人受到电击，但当你认识到自己仅是因为一个身穿实验室工作服的人的命令而去残酷地电击别人（很可能致死），你会作何感想？人们对米尔格拉姆研究方法的批评主要是（参见Baumrind，1964；Miller，1986）在实验期间，被试所承受的压力是让人无法接受的。此外，他们还指出，这种潜在的消极影响可能会长期存在。当被试经历内心折磨后，主试告诉他们这是个骗局，他们会怎样？也许他们会感到自己受人利用，感到难堪，而且有可能在他们未来的生活中不再相信任何心理学家或是合法权威。

另外一种批评意见主要集中于米尔格拉姆研究结果的效度方面（如，Brief et al.，1995；Orne & Holland，1968）。这一批评经常被引用的基础是认为，被试与主试间的关系不仅有信任的成分，还有依赖的成分，并且实验室对于被试来说也是一种陌生的环境，在那里表现出的服从行为并不代表现实生活中的服从行为。因此，批评者声称米尔格拉姆的研究结果不仅效度低，而且由于其效度低下，实验对被试所实施的处理也就是不适当的。

通过在被试参加完实验后对他们进行的调查，米尔格拉姆对这些批评意见作出了回应。他发现84%的被试很高兴参加了这次活动，而只有1%的人对这次经历表现出后悔。另外，一个精神病专家访谈了40名被认为是在实验中感觉最不舒服的被试。他得出的结论是，这些人当中没有一个人遭受长时间的心理折磨。因批评者认为实验室并没有反映现实生活中的行为，米尔格拉姆的意见是：“来到实验室的被试是主动的、有选择权的成人，他有能力接受或拒绝在实验中给予他的命令。”（Milgram，1964，P.852）

这里所报告的米尔格拉姆的研究，引起了人们对以人类为被试的实验中道德问题的深刻思考。事实上，该研究是对心理学服从行为研究领域的影响更大，还是对心理学研究以人类为被试的道德政策问题影响更

大，仍是极具争议的（正如在本书前言中所讨论的）。

近期应用

米尔格拉姆的服从研究继续对目前的研究发挥着广泛的影响，这一点可以在近期研究的不少注解中得到证明。这些研究主要是在米尔格拉姆早期实验方法和结果的启发下进行的。从20世纪60年代早期米尔格拉姆从事他的研究以来，每年都有与此有关的研究出现，这些研究可分为两类：一类是对米尔格拉姆理论的完善以及关于对权威人物服从倾向的详细阐释；另一类是关于在实验中对人类被试施行欺骗手段的道德问题的争论。

作为斯坦利·米尔格拉姆的研究和职业生涯的权威评述者以及米尔格拉姆传记《电醒人心》（The Man Who Shocked the World）的作者，托马斯·布拉斯（Thomas Blass）回顾了源于米尔格拉姆早期服从实验的所有研究及其社会意义（Blass，1999；2002）。总体上来看，布拉斯发现各种研究普遍支持米尔格拉姆最初的研究结果。更重要的是，在米尔格拉姆第一次发表他的结果之后的40多年里，服从率似乎并没有发生显著的改变。这与美国人的直觉判断相矛盾，一般而言，他们会认为自己已经不像40年前的美国人那样敬重权威，当权威人士命令他们做自己不同意的事时，他们会做出反抗。

另一个经常出现的问题是对米尔格拉姆早期研究中性别问题的关注。事实上以前所有的研究对象都是男性。仔细想想，总体而言，男性与女性谁更倾向于服从权威？布拉斯从米尔格拉姆的后期研究和其他人的研究中发现，男人和女人在服从率上并没有什么区别。（想要更详细地了解米尔格拉姆的研究工作所产生的历史影响，请参见布拉斯的网站，网址为：www.stanleymilgram.com。）

与米尔格拉姆的发现紧密衔接的一个研究，调查了路易斯安那州监狱负责执行死刑的“行刑小组”成员的心理体验（Osofsky & Ososky，2002）。研究者访问了50名直接参与行刑的警官，结果发现，尽管比大多数人更多地面对痛苦和死亡，这些被试并没有出现临床上的抑郁症状。被试报告说他们依靠宗教信仰、对同行群体的认同和对责任进行扩散的能力来应对痛苦情感。“然而，这些警官体验到矛盾的情感并且时常报告在执行社会‘终级处罚’时感觉很糟糕。”（P.358）

在伦理方面，一项研究使用米尔格拉姆的研究范式考察在互联网上进行的社会科学研究潜在的棘手的伦理问题（Pittenger，2003）。今天，大量的研究是通过国际互联网进行的，而且此类研究的数量在将来还会显著增加。皮滕杰（Pittenger）主张，研究者必须警惕侵犯隐私、未获得被试的知情同意，使用欺骗手段等潜在的违背道德伦理规范的情况。“互联网为研究者提出了独特的挑战，”皮滕杰写道，“其中包括需要

清楚地界定在网上进行的个人行为与公众行为的区别，确保获得被试有效知情同意的机制以及对被试进行调查询问的机制，核实所收集到的数据的有效性。”（P.45）

一个重要的问题是，我们该如何保护被试在心理学研究中免遭不合理的欺骗手段的伤害，而又同时允许某些科学研究进行所必要的欺骗继续存在？温德勒（Wendler，1996）建议在研究涉及欺骗手段时，研究者应给予被试更高水平的知情同意权（参见本书前言部分对此概念的讨论）。这就意味着，在你同意成为实验被试之前，工作人员会告诉你在此实验中可能会用到欺骗手段，尽管你可能并不会意识到欺骗的确切内容。温德勒声称：“这种对于可接受的欺骗的二级许可给我们提供了最好的机会来协调二者的关系，即既要尊重被试，又要满足科学实验所偶尔需要的欺骗过程。”（P.87）

结论

研究米尔格拉姆的史学家托马斯·布拉斯（2002年）在其所著的关于米尔格拉姆生活和工作的传记中给本研究提供了一个适宜的结论：

> 我们不需要米尔格拉姆告诉我们自己有服从他人命令的倾向。在米尔格拉姆的实验前，我们不知道的是这种倾向的力量有多大。了解我们极强的遵从权威的倾向后，我们可以设法使自己提防不受欢迎或应受谴责的命令（Blass, 2002, P.73）。

Baumrind, D.(1964).Some thoughts on the ethics of research: After reading Milgram’s “Behavioral Study of Obedience.” American Psychologist，19, 421-423.

Blass, T.(1999).The Milgram paradigm after 35 years: Some things we now know about obedience to authority.Journal of Applied Social Psychology, 29(5), 955-978.

Blass, T.(2002).The man who shocked the world.Psychology Today, 35, 68-74.

Blass, T.(2004).The man who shocked the world.New York: Basic Books.

Brief, E., Collins, B., & Miller, A.(1995).Perspectives on obedience to authority: The legacy of the Milgram experiments.The Society for the Psychological Study of Social Issues, 51，1–19.

Milgram, S.(1964).Issues in the study of obedience: A reply to Baumrind.American Psychologist，19, 448-452.

Milgram, S.(1974).Obedience to authority.New York: Harper & Row.

Miller, A.G.(1986).The obedience studies: A case study of controversy in social science.New York: Praeger.

Orne, M.T., & Holland, C.H.(1968).On the ecological validity of laboratory deceptions.International Journal of Psychiatry, 6, 282–293.

Osofsky, M., & Osofsky, H.(2002).The Psychological Experience of Security Officers Who Work with Executions.Psychiatry: International and Biological Processes, 65, 358-370.

Pittenger, D.(2003).Internet research: An opportunity to revisit classic ethical problems in behavioral research.Ethics and Behavior , 13, 45-60.

Wendler, D.(1996).Deception in medical and behavioral research: Is it ever acceptable? Milbank Quarterly, 74(1), 87.